



NFS Update

Calum Mackay

NFS Group

calum.mackay@sun.com



Agenda

- Sharemgr
- In-Kernel Sharetab
- pNFS – Parallel NFS (NFSv4.1)
- NFSv4.0 Namespace extensions
 - > Mirror mounts
 - > Referrals
- NFSv4 over RDMA
- YANFS (formerly WebNFS)

Sharemgr

- Improved file share management (exports)
 - > sharemgr (1m) for managing shares
 - > sharectl (1m) for managing protocol separate from shares
 - > Introduces concept of share groups
 - > Integrates with smf(5)
- NFS support now, CIFS in the future
 - > Extensible via plugin modules
- Available in OpenSolaris since build 53

Sharemgr

- Named share groups hold collection of shares
- Configuration at the group level
 - > common configuration properties
 - > enable/disable by group
- Share level override of properties
- Group “default” for backward compatibility
- Group “zfs” provides handle to ZFS managed shares

Sharemgr

- Each share group is implemented as an SMF *service instance*
 - > instances can start/stop in parallel
- ZFS shares started via SMF but configuration is stored in ZFS properties
 - > dataset with a “*sharenfs*” property appears as sub-group
 - > there is a single “*zfs*” service instance
- Future enhancements include tighter integration with ZFS
 - > ZFS will use sharemgr API

Sharemgr

- Creating a share group and adding shares
 - > sharemgr create homedirs
 - > sharemgr set -P nfs -p nosuid=true homedirs
 - > sharemgr add-share -s /export/home/john homedirs
 - > sharemgr add-share -s /export/home/bill homedirs
 - > sharemgr show -vp homedirs
homedirs nfs=(nosuid=true)
 /export/home/john
 /export/home/bill
 - > svcs group
online 13:00:09 svc:/network/shares/group:zfs
online 13:00:17 svc:/network/shares/group:default
online 14:05:04 svc:/network/shares/group:homedirs

Sharemgr

- Adding security-related settings to “homedirs”
 - > sharemgr set -P nfs -S sys -p rw="*" ro=rohost homedirs
 - > sharemgr show -vp homedirs
homedirs nfs=(nosuid=true) nfs:sys=(rw=* ro=rohost)
 /export/home/john
 /export/home/bill
- Disabling/Enabling a Share Group
 - > sharemgr disable homedirs
 - > sharemgr list -v
default enabled nfs
zfs enabled nfs
homedirs disabled nfs

Sharectl

- Viewing protocol properties
 - > sharectl get nfs
listen_backlog=32
protocol=ALL
servers=16
grace_period=90
nfsmapid_domain=sun.com
server_versmin=2
server_versmax=4
client_versmin=2
client_versmax=4
server_delegation=on
max_connections=-1
- Changing protocol properties
 - > sharectl set -p nfsmapid_domain=sun.com nfs

In-Kernel Sharetab

- Moves `/etc/dfs/sharetab` into memory
 - > Creates a mntfs style filesystem called sharefs
 - > Sharetab maintained (i.e. updated) solely in the kernel
 - > `/etc/dfs/sharetab` still appears as a read-only file
 - > Improves start up times for machines with large number of shares
- OpenSolaris: available now, since build 62

pNFS – Parallel NFS

- Clients are given the ability to access the data servers directly
- The data can be striped across many data servers
- Backwards compatible
 - > Can access data via the metadata server
- Standards-based: heterogeneous support

pNFS



NFS Client

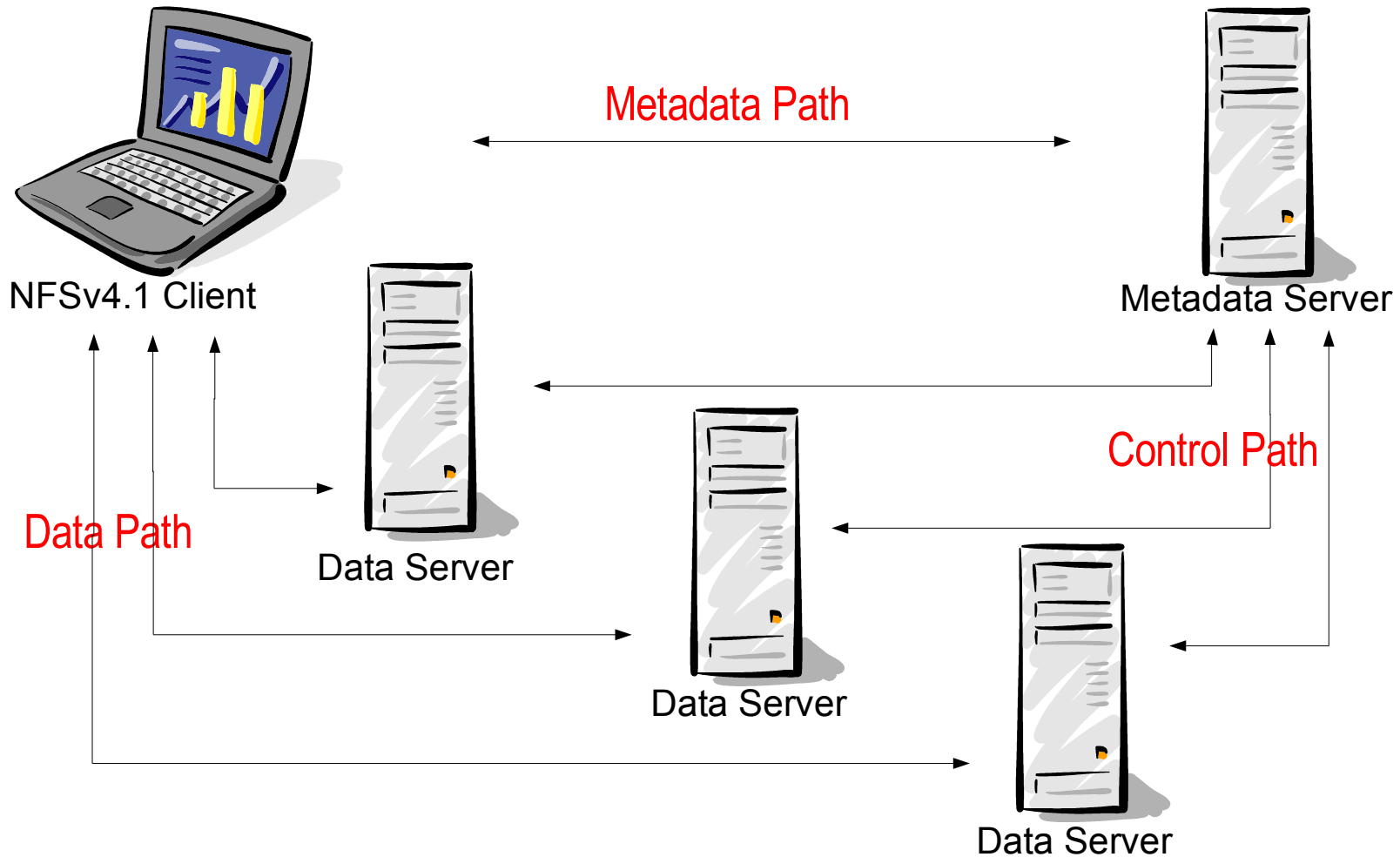
Data and Metadata



NFS Server

Today...

pNFS



pNFS

- High Performance
 - > Provide scalable parallel access to files distributed among multiple data servers
- Horizontal Scalability
 - > Reduce the resource limitation of the single NFS server serving all of the files in an exported file system
- Separation of Data and Metadata
 - > easy to move data around
 - > dynamically notify client of data location

pNFS

- Implementation on OpenSolaris
 - > Currently in prototype phase, rapidly firming up product plans
 - > OpenSolaris project launched late last year
 - > <http://www.opensolaris.org/os/project/nfsv41>
- Implementations in the Industry
 - > Also in the prototype phase ..
 - > most NFS vendors
- The protocol is still relatively new
- A lot of interoperability testing going on

Mirror-mounts

- Enabling an NFSv4 client to traverse server filesystem boundaries
 - > As the client discovers new shared filesystems on the server, it automatically mounts them
 - > Eliminates the need for automounter maps and the associated administrative overhead
- Similar to automounter maps
 - > but automatic
- Handy when the server shares hundreds or thousands of ZFS filesystems

Mirror-mounts

- Implementation on Solaris/OpenSolaris
 - > Prototype complete, requirements/design underway
 - > OpenSolaris Project: NFSv4 namespace extensions
 - <http://www.opensolaris.org/os/project/nfs-namespace/>
 - > We are actively looking for participants from the community ..

Referrals

server-1



/export/not_here

server-2



/export/here



A mechanism by which a server can “point” to another server

Referrals

- Underpinning technology to enable multi-server namespace construction
- Similar to automounter maps
- Part of the NFSv4.0 specification (fs_locations attribute)
- Plus extensions to NFSv4.1 protocol (fs_locations_info, fs_absent attributes)
- Futures include Replication and Migration

Referrals



Give me "foo"

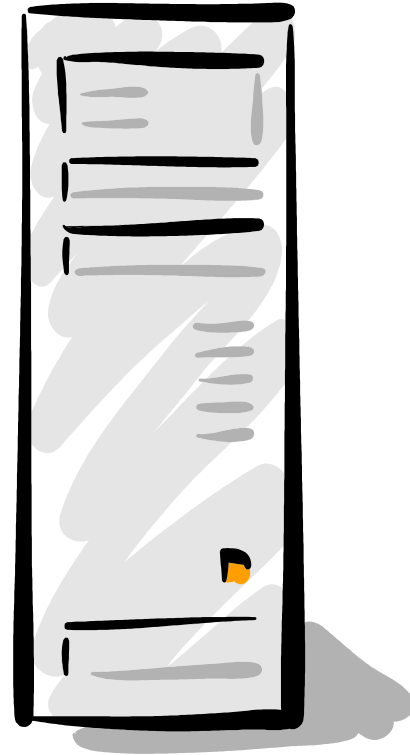
It's not here

NFS4ERR_MOVED

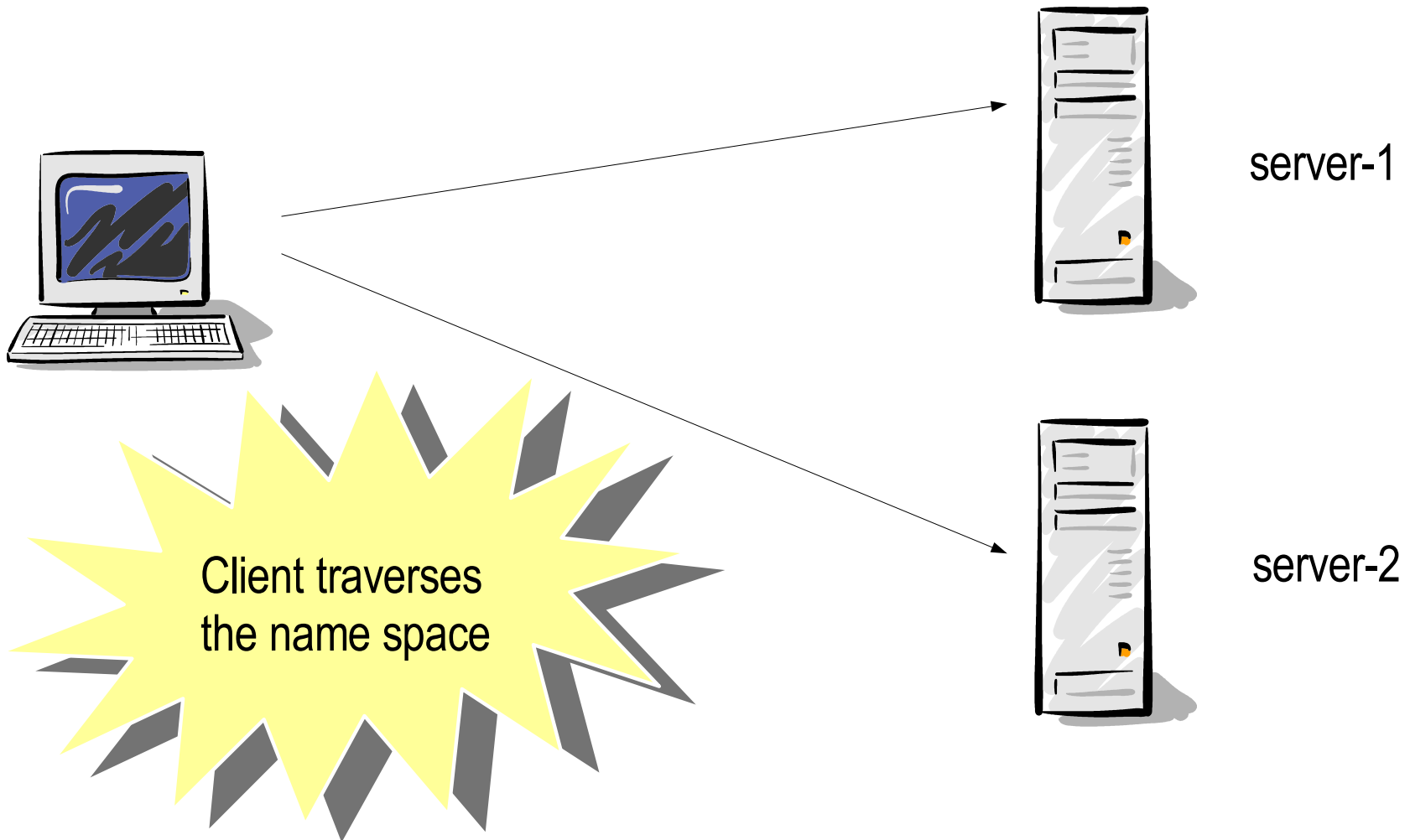
Where?

Over there

fs_locations



Referrals



Referrals

- Implementation on Solaris/OpenSolaris:
 - > Prototype complete, requirements/design underway
 - > Initial support to include only client side v4.0 referrals
 - > OpenSolaris Project: NFSv4 namespace extensions
 - <http://www.opensolaris.org/os/project/nfs-namespace/>
 - > We are actively looking for participants from the community ..

NFS over RDMA

- Update on Solaris/OpenSolaris code:
 - > Work done by Ohio State University to improve NFS/RDMA performance and bring the code up to spec
 - > Working on a plan to get the code changes back to OpenSolaris
 - > <http://opensolaris.org/os/project/nfsrdma/>
 - > Additional year of funding to OSU to continue NFS/RDMA work in the context of interoperability and pNFS

YANFS

- YANFS is the new WebNFS
 - > A project to provide NFSv4.1 client & server, in Java
 - > open source
 - > In early stages...
 - > <http://yanfs.dev.java.net/>
 - > <http://blogs.sun.com/shepler/>
 - > we're looking for contributors



NFS Update

Calum Mackay

NFS Group

calum.mackay@sun.com

