Managing an OpenSolaris pNFS Server

Lisa Week

24 February 2009

Topics

- Identify problems with managing pNFS servers (MDSes and DSes)

- Solving these problems - OpenSolaris pNFS Server Management Model

Problems

- Security
  - > How do you prevent rogue data servers from entering the community?
  - > How do you make sure admins/users can't cause Data and Metadata to get out of sync?

- Ease of Use
  - > How do you make it easy to manage a community with a large number of machines?
  - > How do you control where your files go?

- Observability
  - > How do you know when things are going well? How do you know when things are going wrong? How do you diagnose problems?
  - > How do you find out which data servers hold data for a file system?

MDS Setup

- Commands to set up a metadata server

  (Note: modifications to existing commands are in blue)

  > `#zpool create mdspool /dev/dsk/c0t0d0s7`

  > `#zfs create -o pnfs=on,sharenfs=on mdspool/pnfs`

    > `pnfs=on` indicates the type file system is being used as a pNFS metadata server

    > `sharenfs=on` indicates whether the file system is shared or not

DS Setup

- ## Commands to set up a data server

 (Note: modifications to existing commands are in blue)

  > `#zpool create dspool /dev/dsk/c0t0d0s7`

  > `#zfs create -o mds=192.168.1.1,sharepnfs=on -t pnfsdata dspool/pnfs`
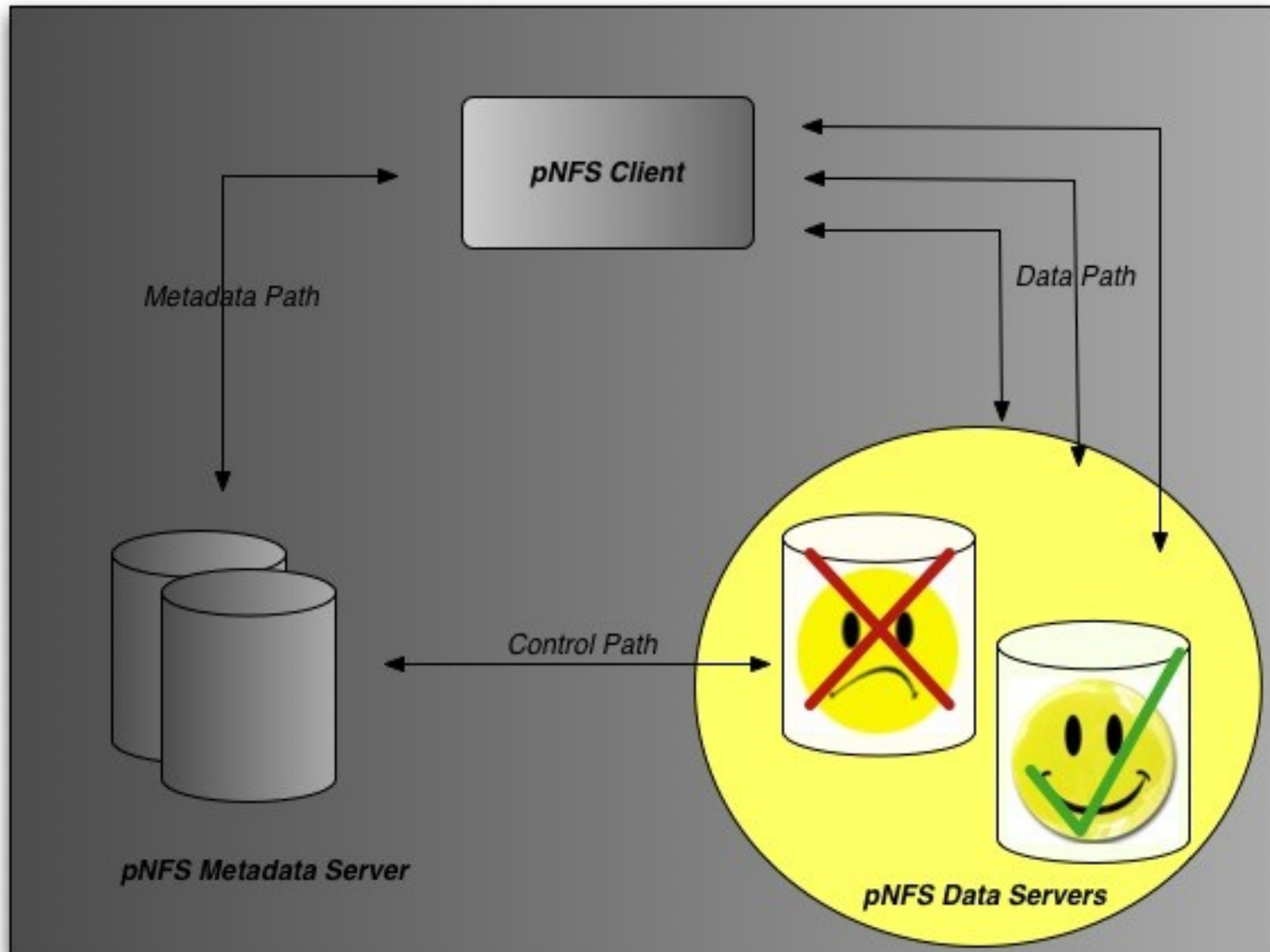
    > `-t pnfsdata` indicates the type of dataset being created

    > `mds` property indicates the Metadata Server for this dataset

    > `sharepnfs` property indicates whether the dataset is shared or not (i.e. whether or not the `mds` will use it for I/O)
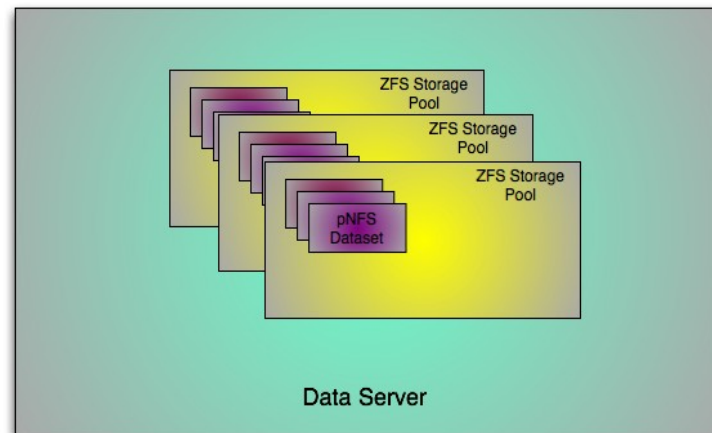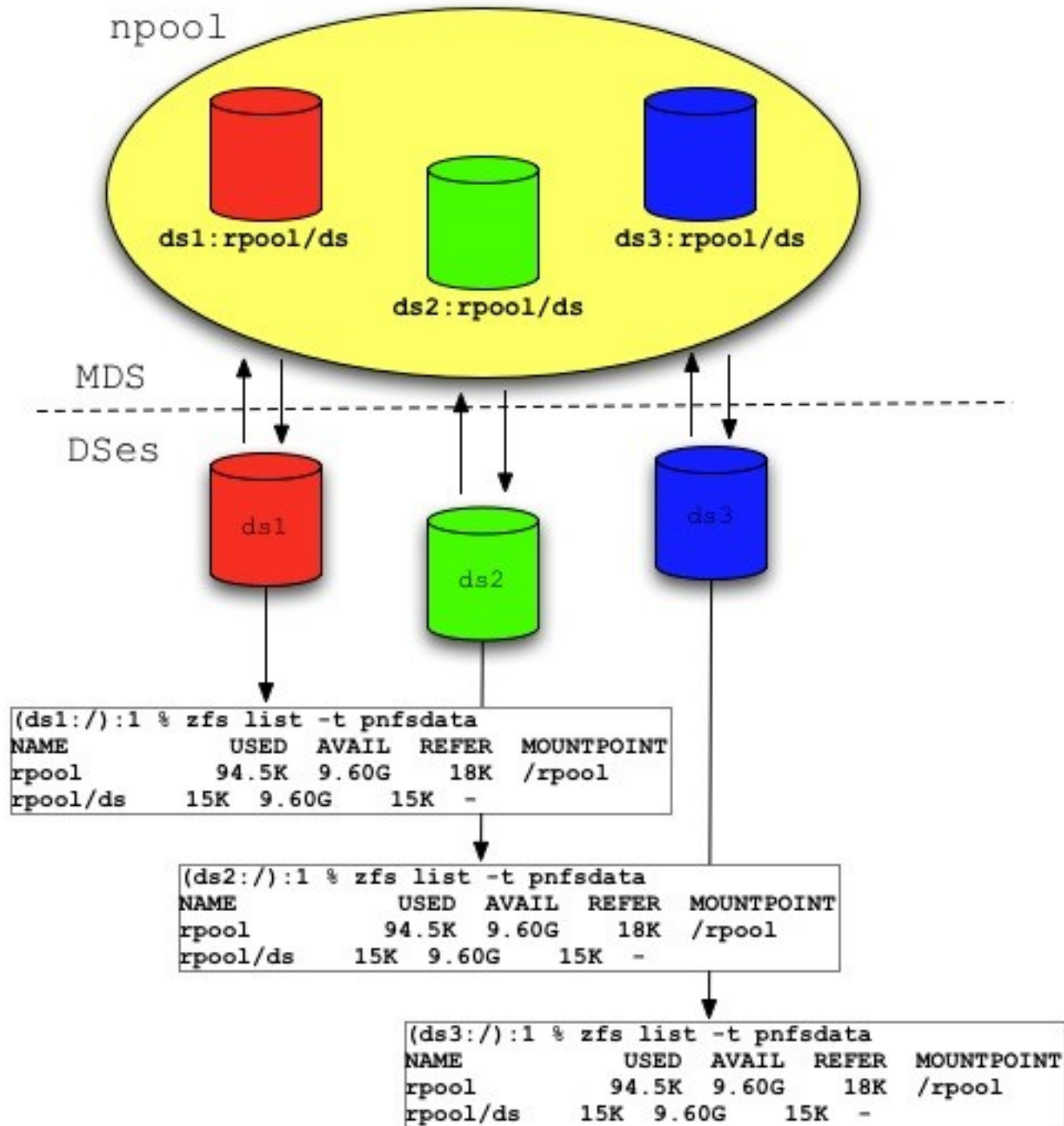
Security: No rogue data servers...

MDS Setup

- New command called `nfsadm` used to:
  - > *Accept* or *reject* data servers and data sets
  - > Manage *npools*
  - > Manage *policies*

- *npools* = "network pools"; objects used to put data server data sets into manageable groups at the MDS

npools

MDS: Simple Policy Engine

- Policies are "create-time" only

- Policies define parameters; path, extension, owner and group, time and date

- Policies can be set at client and MDS
  - > Those set at the client are "hints" and only describe striping parameters (e.g. 4 way stripe, 32K interlace). Additionally, the server may ignore the hint (e.g. It has a conflicting policy defined)
  - > Those set at the server describe striping parameters and *npools* to use

How do I find out what is going on?

- Policy "explain" mechanism
- DTrace providers on the server (NFSv3,4.0, 4.1)
- Extend snoop/wireshark to talk the control protocol
- nfsstat -s
- nfsstat -l (to tell the layout of the file) is not available on the server

Server Management Futures

- Centralized administration
- Higher level and more centralized monitoring of the pNFS servers

Manging an OpenSolaris pNFS Server

lisa.week@sun.com