

# Advanced Unix System Administration

Lecture 21  
April 18, 2007

Steven Luo  
<sluo+decal@OCF.Berkeley.EDU>

# Network Storage

- Two forms of network storage:
  - Disk-level access – give each client access to the underlying physical disk
    - Clients do reads from the disk and have to understand the filesystem themselves
    - Server can be simple, but more network I/O
    - SAN implementations (Fiber Channel, iSCSI)
  - File-level access – clients operate on files
    - Clients request files, server fulfills requests
    - Server is more complex, saves on network I/O
    - NFS, SMB/CIFS, AFS/Coda ...

# Network Storage

- The Network File System (NFS)
  - First developed at Sun in the 1980s, now controlled by IETF
  - Versions 2 and 3 attempt to provide (mostly) stateless operation, to simplify crash recovery
    - Features requiring state (locking) are implemented in additional protocols
  - Attempts to provide POSIX-like semantics within the possibilities of stateless operation
    - Note that NFS is NOT a POSIX-compliant filesystem!

# Network Storage

- The Network File System (NFS) con't
  - Lifecycle of an NFS request
    - Client requests a “file handle” from the NFS mountd
    - Client requests that the nfsd perform operations on this file handle
      - Operations should be simple, atomic, and idempotent
      - In reality, they're not always
    - nfsd returns only when it has finished the operation or the operation has failed

# Network Storage

- The Network File System (NFS) con't
  - Problems (in traditional NFS)
    - It's difficult to guarantee that repeating operations will be safe
    - Some features of POSIX filesystems, especially those requiring atomicity, are impossible to implement
    - The default mode can leave clients hanging for a very, very long time on a server crash
    - Security depends on the client saying who it is

# Network Storage

- Other network file systems
  - As a rule, all maintain state
    - Allows for more complex commands, more featureful semantics, but makes crash recovery difficult
  - Most implement richer access control
    - The exact model tends to differ – AFS, CIFS, and NFSv4 offer separate, differing access control models
    - Over-the-wire security may also be present
  - Some provide for distributed operation

# Network Storage

- iSCSI
  - A translation of the SCSI command set to a network
    - SCSI topology already looks a bit like a network, so this isn't too hard
    - Reliability questions need to be dealt with
  - An “initiator” on the host connects to storage on a “target”
  - Simultaneous access from multiple initiators requires filesystem support (Red Hat GFS, OCFS)