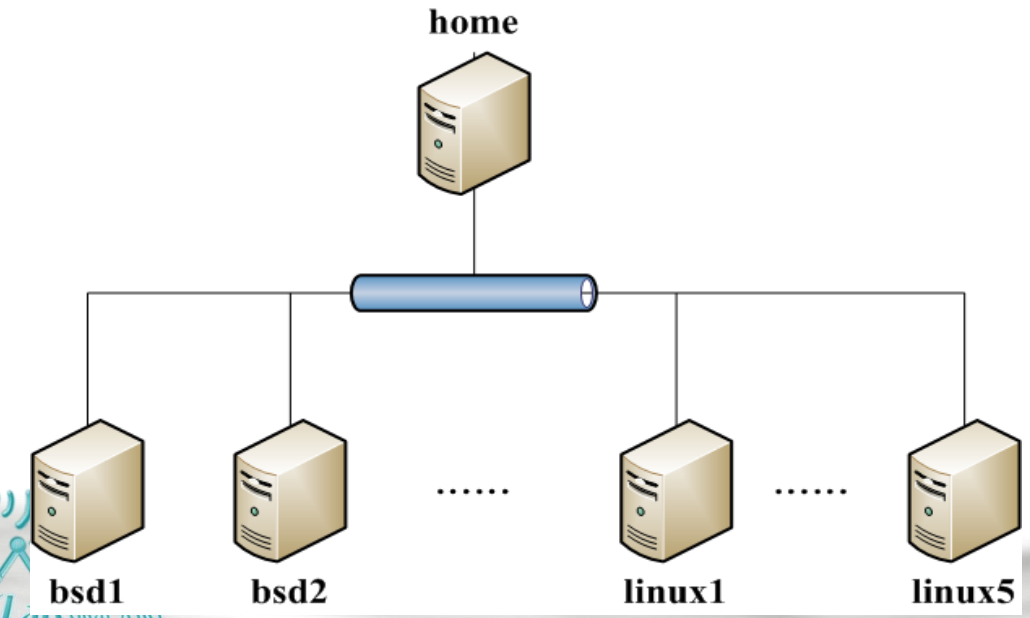
Network File System

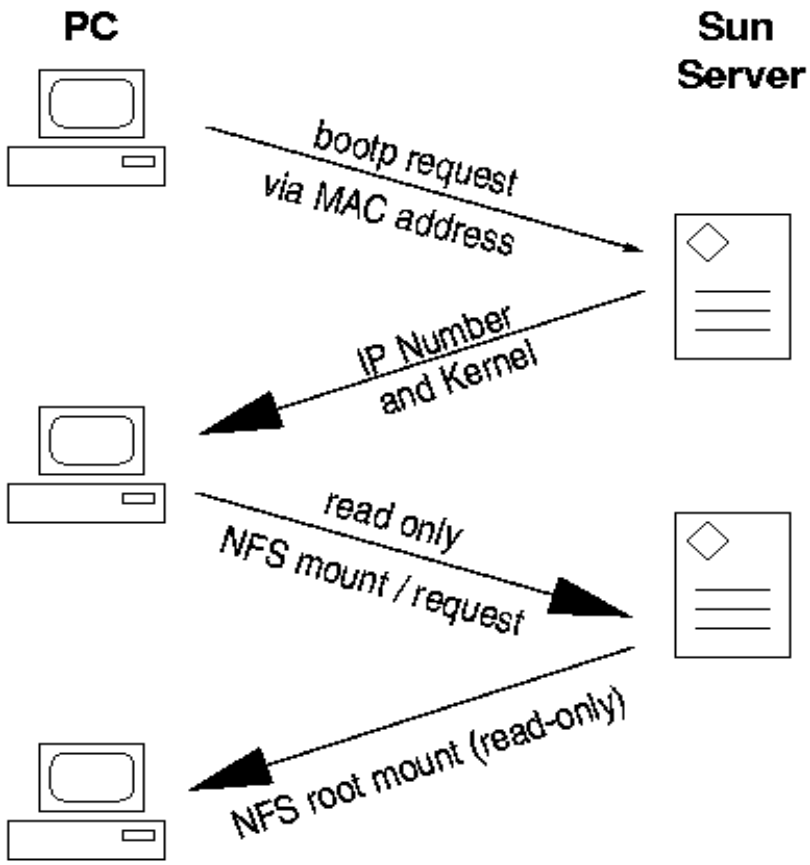
Computer System and Network Administration

Department of Computer Science & Information Engineering
National Cheng Kung University
2016 Fall

NFS

- Share filesystem to other hosts via network
- NFS History
  - Introduced by Sun Microsystems in 1985
  - Originally designed for diskless client-server architecture





The PC then starts the appropriate X-Server using the MAC address as a key

Components of NFS

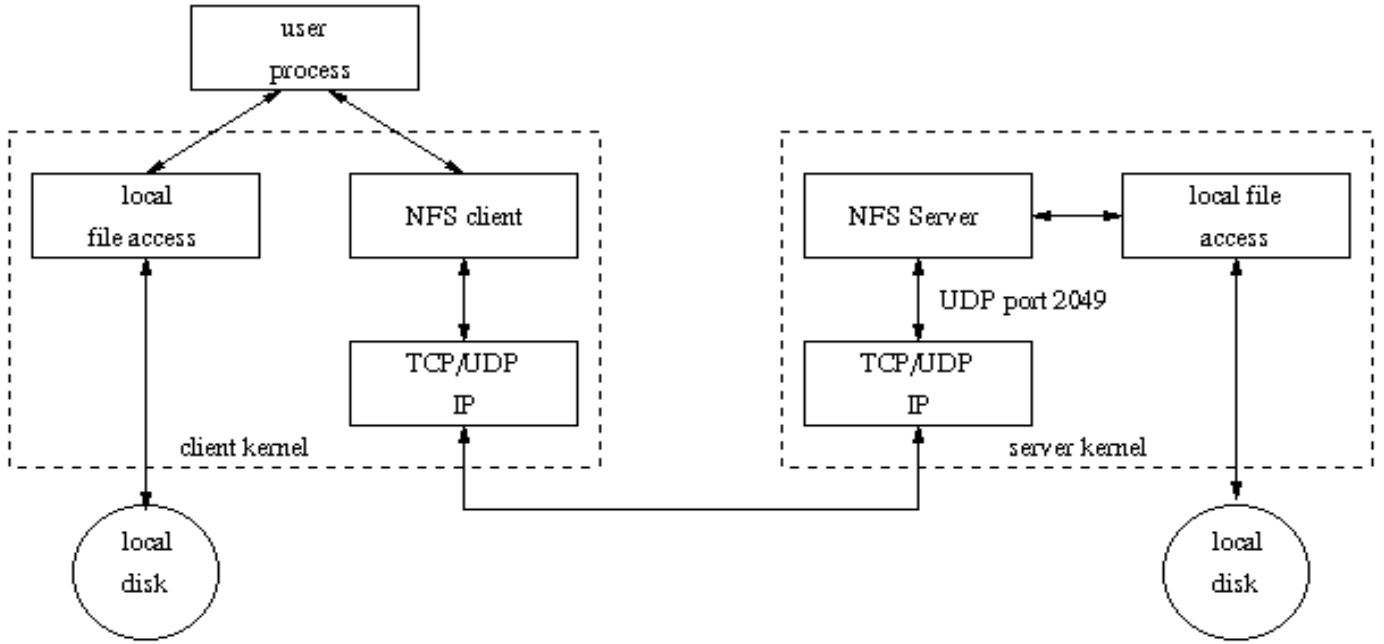
- Including
  - Mounting Protocol
  - Mount Server
  - Daemons that coordinate basic file service
  - Diagnostic utilities

Components of NFS – mounting protocol (1)

- NFSv2
  - Synchronous write
    - V2 NFS server must commit each modified block to disk before replying to NFS client
    - Cause long delay when there is a NFS write operation
- NFSv3 in 1990s
  - Asynchronous write
    - Provide increase performance and better support for large files
- NFSv4 in 2000s
  - Available in FreeBSD 8.1-R
  - Stateful protocol
  - Unicode support

Components of NFS – mounting protocol (2)

- Sun's ONC distributed computing standards
  - NFS client → RPC → Transport Layer → …
  - Transport Layer
    - UDP: Lack congestion control
    - TCP: become more suitable



Components of NFS – mounting protocol (3)

- Advanced NFS feature support by OS

System	NFSv3	TCP	Default
FreeBSD	Yes	Yes	UDP
Linux (debian)	Yes	Yes	UDP
Solaris	Yes	Yes	TCP

Components of NFS – Server-side NFS (1)

- NFS Server
  - Export sharing filesystem
    - System dependent
  - Waiting for “mount request”
    - mountd (rpc.mountd) daemon
  - Waiting for “file access request”
    - nfsd (rpc.nfsd) daemon

Components of NFS – Server-side NFS (2)

- Exporting filesystem
  1. Edit export configuration file
     - Each line is "what to export and how"
  2. Reload related daemons

System	Exports info file	How to reload
FreeBSD	/etc/exports	/etc/rc.d/mountd restart
Linux	/etc/exports	/usr/sbin/exportfs -a
Solaris	/etc/dfs/dfstab	/usr/sbin/shareall
SunOS	/etc/exports	/usr/sbin/exportfs -a

You can also reload by using "kill -1 <mountd's pid>"

Components of NFS – Server-side NFS (FreeBSD.1)

- Exporting filesystem
  - /etc/exports
    - White-space separated
    - Format: *directory-list options-list client-list*

Option	Description
-ro	Exports read-only, default is (read-write)
-alldirs	Allow any subdirectory to be mounted
-maproot=user	Maps root to the specified user.
-mapall=user	Maps all UIDs to the specified user.

Client	Description
hostname	Host name (ex: mailgate ccserv)
netgroup	NIS netgroups
-network -mask	-network 140.116.235.0 -mask 255.255.255.0

Components of NFS – Server-side NFS (FreeBSD.2)

- Example of /etc/exports

```
/raid       -alldirs  -maproot=root mailgate ccserv backup
/raid       -alldirs  -maproot=65534  -network 140.116.209.0  -mask 255.255.255.0
/home       -ro  -mapall=nobody  -network 140.116.235.0  -mask 255.255.255.0
/usr/src  /usr/obj  -maproot=0 bsd_cc_csie
```

  - Network and mask cannot appear on the same line with hosts and netgroups

- Reload daemons
  - % kill -1 `cat /var/run/mountd.pid`
  - /etc/rc.d/mountd restart

Components of NFS – Server-side NFS (Linux.1)

- Exporting filesystem
  - /etc/exports
    - Format: *directory client-list-with-option*
    - Ex: /home1 ccbsd5(ro)

Client	Description
hostname	Host name (ex: mailgate ccserv)
@netgroup	NIS netgroups
ipaddr/mask	CIDR-style specification (ex: 140.116.235.0/24)
Wild cards * ?	FQDN with wild cards (ex: bsd*.csie.ncku.edu.tw)

Components of NFS – Server-side NFS (Linux.2)

Option	Description
ro,rw	Read-only, Read-write (default)
rw=list	Hosts in the list can do rw, others ro only
root_squash	Maps UID 0 and GID 0 to the value of anonuid and anongid (default)
no_root_squash	Allow root access
all_squash	Maps all UID and GID to anonymous one
subtree_check	Check that the accessed file is in the appropriate filesystem and in the exported tree.
no_subtree_check	Disables subtree checking
anonuid=xxx	Related to root_squash
anongid=xxx	Related to root_squash
secure	Require remote access from privileged port
insecure	Allow remote access from any port
noaccess	Prevent access to this dir and its subdir

Components of NFS – Server-side NFS (Linux.3)

- Example of /etc/exports

```
/home1          sun*.csie.ncku.edu.tw(rw)
/home2          @sun_cc_csie(ro)  dragon(rw,no_root_squash)
/home           ccpc1(rw,all_squash,anonuid=150,anongid=100)
/ftp/pub        (ro,insecure,all_squash)
/users          *.xor.com(rw)
/users/evi      (noaccess)
```

- Run /usr/sbin/exportfs
  - % /usr/sbin/exportfs –a
    - Maintain /var/lib/nfs/xtab table which is read by mountd

Components of NFS – Server-side NFS (Solaris.1)

- Exporting filesystem
  - /etc/dfs/dfstab
  - Each line will execute "share" command to export one NFS
    - [format] share –F nfs –o option-list directory
    - Ex: share –F nfs –o rw=bsd5.csie.ncku.edu.tw /home2
- Run shareall command
  - % /usr/sbin/shareall

Client	Description
hostname	Host name (ex: mailgate ccserv)
netgroup	NIS netgroups
IP networks	@CIDR-style specification (ex: @140.116.235.0/24)
DNS domains	.xxx.yyy any host within the domain (ex: .ncku.edu.tw)

Components of NFS – Server-side NFS (Solaris.2)

Option	Description
ro,rw	Read-only to all, Read-write to all
ro=list, rw=list	Hosts in the list can do ro/rw
root=list	Lists hosts permitted to access this filesystem as root. Otherwise, root access from a client is equivalent to by "nobody"
anon=xxx	Specify the UID to which root is remapped. Default is "nobody"
anongid=xxx	Related to root_squash
nosub	Forbids clients to mount subdirectories
nosuid	Prevents setuid and setgid from being created

Components of NFS – Server-side NFS (3)

- nfsd daemon
  - Handle NFS file access request from NFS clients
  - Number of nfsd is important
    - Too small, some NFS request may not be served
    - Too large, load will be high
- In FreeBSD
  - Specify nfsd options in /etc/rc.conf
    - nfs_server_enable="YES"
    - nfs_server_flags="-u –t –n 4"

Components of NFS – client-side NFS (1)

- NFS Client
  - Mount NFS filesystem first
  - Access file under NFS filesystem
- mount command
  - *mount* [-o options] *host:directory mount-point*
  - *Ex.* **% mount –t nfs bsd4:/home/www /home/nfs/www**
- /etc/fstab (/etc/vfstab in Solaris)
  - **% mount –a –t nfs (FreeBSD, Linux)**
  - **% mount –a –F nfs (Solaris)**

```
# Device              Mountpoint     FStype  Options      Dump  Pass#
dragon:/usr/man       /usr/man  nfs  ro,bg,soft  0       0
ccserv:/spool/mail    /var/mail  nfs  rw,bg,intr  0       0
```

  - **Note that the *soft* option may abort programs (e.g., simulation) after running for several hours due to transient network glitch.**

Components of NFS – client-side NFS (2)

- NFS mount flags

Flag	Systems	Description
ro or rw	S,L,F	Mount the NFS as ro or rw
bg	S,L,F	If failed, keep trying in background
hard	S,L	If server down, access will keep trying until server comes back
soft	S,L,F	If server down, let access fail and return error
intr, nointr	S,L,F	Allow/Disallow user to interrupt blocked access
retrans=n	S,L,F	# of times to repeat a request before error return
timeo=n	S,L,F	Timeout period of requests (tens of seconds)
rsize=n	S,L,F	Set read buffer size to n bytes
wsize=n	S,L,F	Set write buffer size to n bytes
vers=n	S	Selects NFS v2 or v3
nfsv3,nfsv2	F	Selects NFS v2 or v3
proto=prot	S	tcp or udp
tcp	L,F	Select TCP. UDP is default

Components of NFS – client-side NFS (3)

- Client side daemons that enhance performance
  - nfsiod (local NFS asynchronous I/O server) or biod (block I/O daemon)
  - Perform read-ahead and write-behind caching

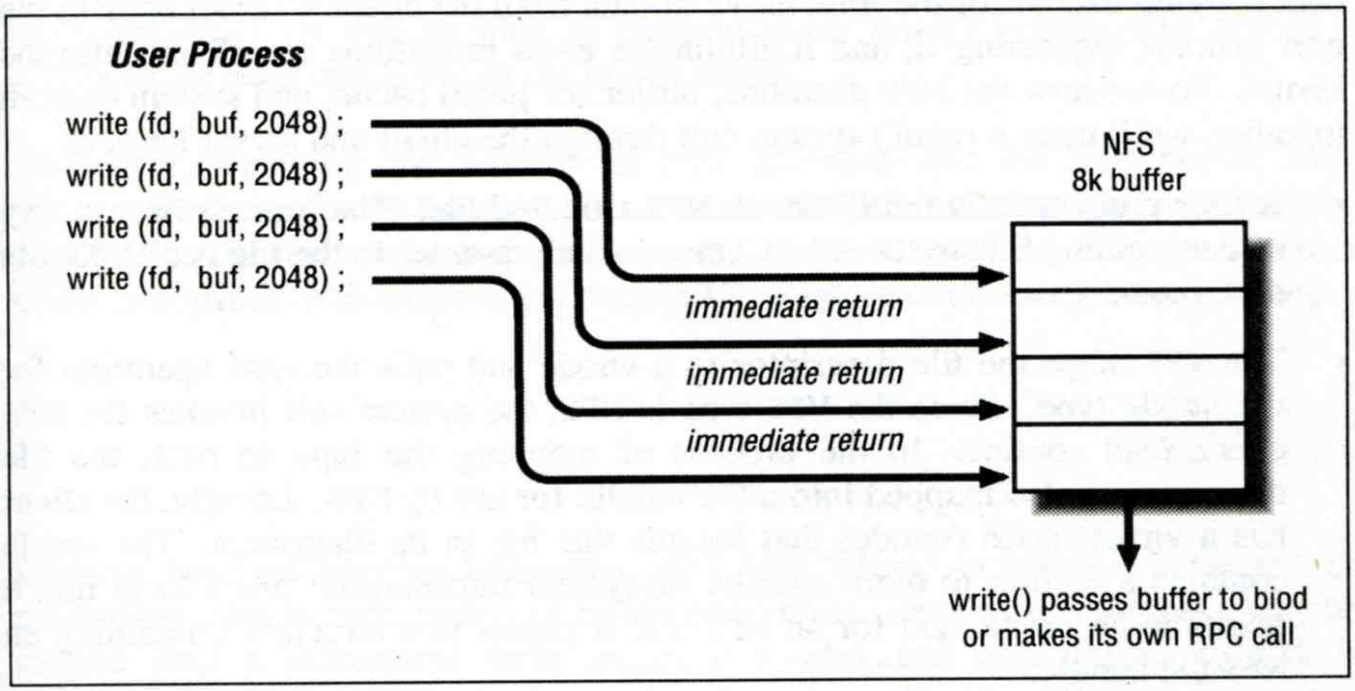


Figure 6-2. NFS buffer writing

Components of NFS – NFS Utilities (1)

- nfsstat - Display NFS statistics
  - % **nfsstat –s** (display statistics of NFS server)
  - % **nfsstat –c** (display statistics of NFS client)

```
> nfsstat -c
Client Info:
Rpc Counts:
   Getattr    Setattr    Lookup  Readlink       Read      Write     Create     Remove
   1065253      34196    379742      5187     111699     182603      18049      29803
    Rename       Link   Symlink     Mkdir      Rmdir    Readdir   RdirPlus     Access
     20838       4746         1        10       1003       4705          0     316560
     Mknod     Fsstat    Fsinfo  PathConf     Commit
         0      13742      3889         0      75747
Rpc Info:
 TimedOut    Invalid  X Replies    Retries   Requests
        0          0         69       3994    2267773
Cache Info:
  Attr Hits     Misses  Lkup Hits     Misses  BioR Hits     Misses  BioW Hits     Misses
    1920497    1259363    1256973     379714     352854     102015     521158     182603
BioRLHits     Misses  BioD Hits     Misses  DirE Hits     Misses
     347749       5187      14996       4685       6137          0
```

Components of NFS – NFS Utilities (2)

- showmount
  - % showmount –e cchome
    - show the hosts's export list
  - % showmount –a
    - List all mount points

```
cshome> showmount -a
All mount points on localhost:
bsd1:/home2
bsd1:/raid/home
csduty:/home2
csduty:/raid/home
linux1:/raid/home
linux2:/raid/home
nat235.dynamic:/raid/home
sun1:/raid/home
```

```
magpie> showmount -e magpie
Exports list on magpie:
/home                ccduty mailgate 140.116.209.0
/drongo              operator ccduty mailgate 140.116.209.0
```

NFS in FreeBSD

- NFS server
  - Edit /etc/rc.conf

```
…
nfs_server_enable="YES"
nfs_server_flags="-u –t –n 4"
…
```

- NFS client

```
…
nfs_client_enable="YES"
…
```

Reference

- RFC 1094 - NFS: Network File System Protocol specification (NFSv2) http://www.faqs.org/rfcs/rfc1094.html
- RFC 1813 - NFS Version 3 Protocol Specification (RFC1813) http://www.faqs.org/rfcs/rfc1813.html
- RFC 7530 - Network File System (NFS) Version 4 Protocol http://www.faqs.org/rfcs/rfc7530.html
- Sandberg, Russel, et al. "Design and implementation of the Sun network filesystem." *Proceedings of the Summer USENIX conference*. 1985.
- FreeBSD Handbook - 28.3 Network File System (NFS) https://www.freebsd.org/doc/handbook/network-nfs.html
- Linux NFS-HOWTO http://www.tldp.org/HOWTO/NFS-HOWTO/index.html