



Defending Against Broken NFS Clients

John Corbin

President

EP Network Storage Performance
Lab, Inc.

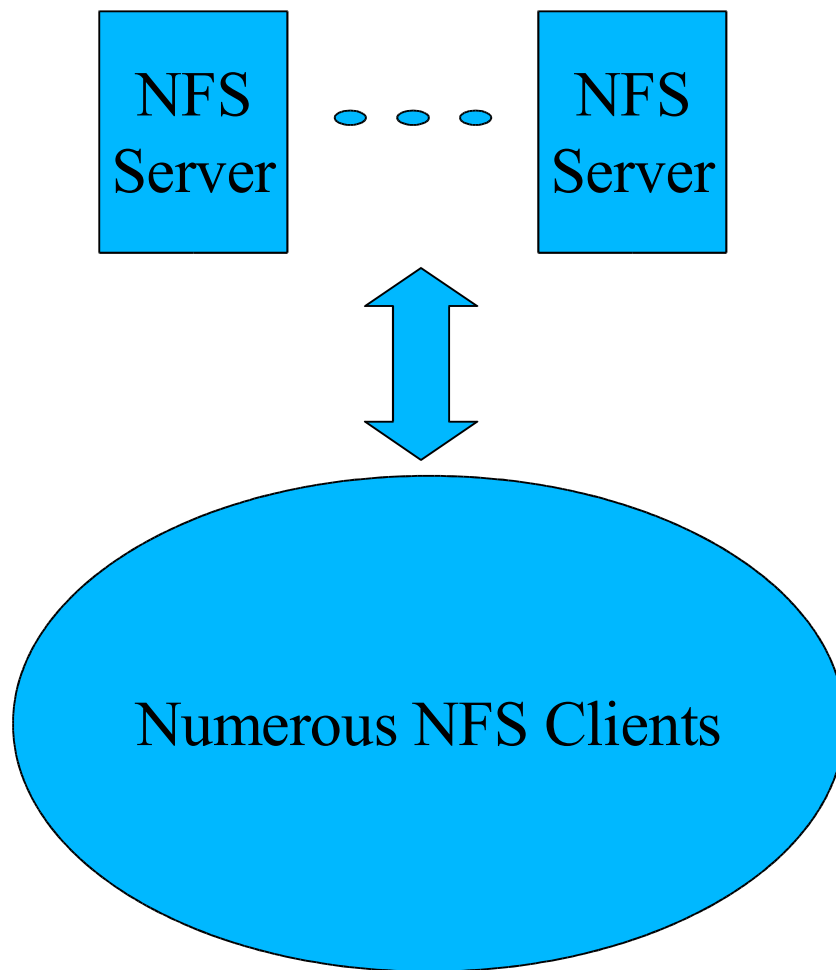
jcorbin@nsplab.com



Overview

- Server/Client Configuration
- Broken NFS Clients
- NFS Server Issues
- Other Potential Issues

Configuration



Configuration Reliability

- Servers should just work
- Clients
 - 5000 Clients
 - Assume 500,000hr MTBF (HW+SW)
 - Failure every 100 hours?
 - Improve MTBF by regularly rebooting
 - Not all at once please

Broken Clients

- Send invalid parameters
 - Make directory inside a file
 - Bad name length
 - Corrupt file handle
- Why?
 - Quality of NFS client implementations
 - Getting Better
 - Usually due to kernel corruption
 - V4 Clients – new and improving

Example

**** Mount/Write File/Unmount Ctrl-C - restart script ****

mount: nfssrv:/testing already mounted or /t/dir busy

mount: according to mtab, nfssrv:/testing is already mounted on /t/dir

doit: line 13: 22530 Segmentation fault umount /t/dir

**** Running mount/umount in a loop ****

mount: wrong fs type, bad option, bad superblock on nfssrv:/testing, or
too many mounted file systems

Broken Client Impact

- Cause repetitive panics on servers
 - Clients should be using hard mounts
 - Causes requests to be retransmitted until the server sends a reply
 - The retransmitted bad request will keep crashing the server
 - Can eventually take down all nodes on a server cluster
- Unintentional DOS Attack

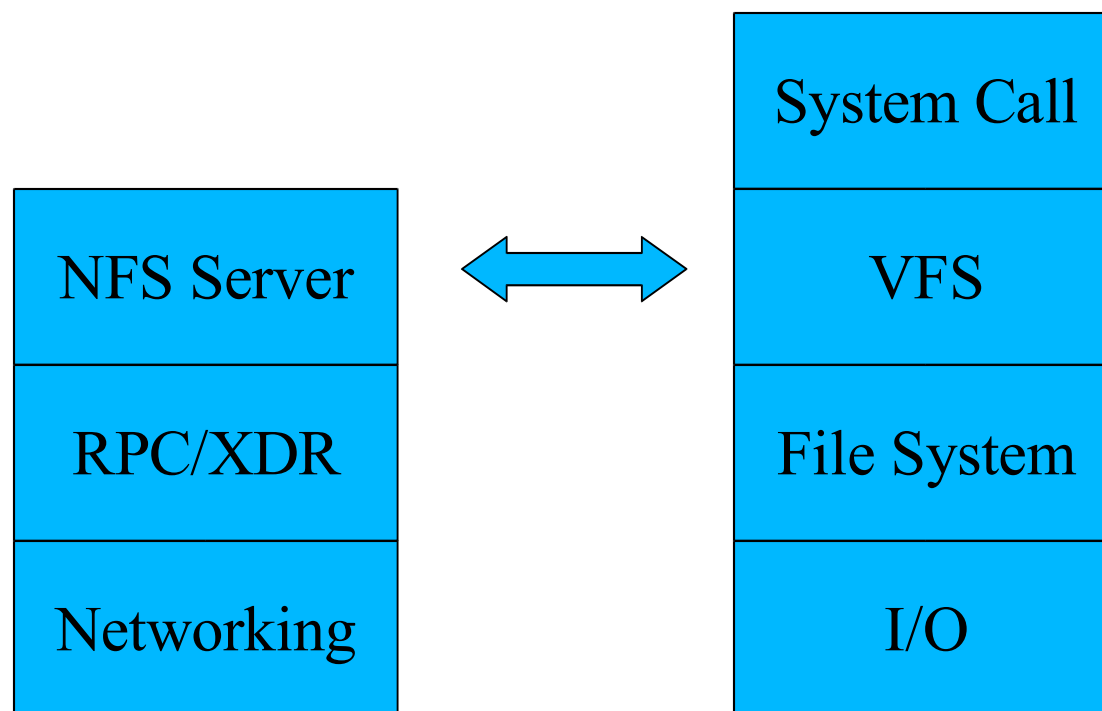
Server Issues

- Weak on request validation
- Handling bad parameters
 - Not the end of the world
 - Don't panic
- Some issues tied to NFS server heritage
- Need for NAS certification product

Traditional Server Architectural Issues

User Level

Kernel Level



Server Arch Issues

- Originally designed around local access
 - System call side validated parameters
 - Panic in file system if parameters bad
- NFS Server is network interface to VFS
 - Bad parameters not a result of locally corrupted kernel
 - Client has the issue

Handling Broken Clients

- Are we crashing in the NFS server software?
 - Customer support can help here
- Turn off NFS service
- Turn on network tracing
 - This helps later
- Reboot broken client

Which Client?

- Find out which client is causing the problem and reboot it
 - Easier said than done
 - Vendor tools ???
 - Get client IP address from the NFS request that caused the failure
 - Usually requires kernel debugging expertise

Which Client?

- IP Trace
 - Shouldn't last packet in trace be the bad one?
 - Not if you have traced from the server
 - Good chance otherwise
 - Assuming you can tell last packet by timestamp
 - Difficult with a lot of clients active
- Need to narrow the problem down

Process of Elimination

- Create export files with partitioning
 - Divide 5000 clients into 5 groups
 - Enable each group until server crashes
 - Let each group run for at least 5 minutes
 - Client has slowed down retransmission rate
 - Divide each group into 5 sub-groups
 - Enable each sub-group till server crashes
 - Narrowed it down to 200 machines
 - Network trace should be useful

Other Potential Issues

- Mount Storms
 - V2/V3 thing
 - Usually caused by automounter
- portmapper
 - Hammered during mount storms
- Reliable Naming Services
 - V4 passes around names not IDs
 - What happens when NFS server cannot resolve names?

Conclusion

- **System Administrators**
 - Pre-build export files with partitioning
 - Request diagnostic tools from vendors
 - Request NAS Certification
- **Server Vendors**
 - Improve NFS request validation
 - Don't panic
 - Deliver diagnostic tools
 - Certify NFS Servers



EP Network Storage Performance Lab, Inc.

“We don’t make NAS and SAN
devices, we make them better.”

www.nsplab.com