NFS with Linux: Current and Future Efforts

Chuck Lever, Network Appliance, Inc
Steve Dickson, Red Hat
Red Hat Summit 2006
Overview

- Linux NFS: Present
- Linux NFS: The Future
- Deploying Linux NFS
- Open Discussion
Linux NFS: Present

Steve Dickson, *Red Hat*

Red Hat Summit 2006
Agenda

- Recent RHEL4 Improvements
- Secure NFS Explained
- Basic NFS V4 Functionality
RHEL4 Improvements

- Database improvements
  - Async I/O with Direct I/O
- Caching improvements
  - Memory mapped files
  - Invalidations
- Mounting improvements
  - More automounts at once
  - Use UDP first, then TCP.
RHEL4 Improvements (continued)

- Coherency
  - Better Close to Open on coherency.
  - -o nocto to turn off on mostly read-only mounts

- Better SMP locking.
  - Attempts to eliminate of the Big Kernel Lock (BKL)

- NFS v3 POSIX ACL support
  - ACL Cache
  - -o noacl turns of all ACL processing
Secure NFS Explained

- Used by ALL three NFS versions
  - Use the ‘-o sec=krb5’ mount option
- Uses GSS-API cryptographic method.
- Three Kerberos 5 security levels
  - Authentication (RPC header is signed)
  - Integrity (Header and Body are signed)
  - Privacy (Header signed. Body encrypted)
Secure NFS (cont'd)

- User level daemons used to handle complicated context initiation phase
  - rpc.gssd – Client daemon that handles security contexts
  - rpc.svccgssd – Server daemon that handles security contexts
- Set SECURE_NFS in /etc/sysconfig/nfs
- Both daemons use files in the rpc_pipefs filesystem to get “upcalls” from the kernel.
Security Context Data flow

- Security Context Needed
- None cached; upcall to rpcgssd
- Server called; upcall to rpcsvcgssd

- rpcsvcgssd does gssapi magic
- Server returns gss context
- gss context cached in client
Basic NFS V4 functionality

- Compound Procedures
  - Multiple operations sent in one Over-The-Wire message.

- Firewall Friendlier
  - Mount and locking protocols are integrated into protocol
  - Only TCP is supported

- Open and Close Operations
  - Atomic creates supported
Basic NFS V4 functionality

- Pseudo File System
  - Shared server namespace
- ID mapping
  - "name@domain" strings are mapped to user id (i.e. integers) by the rpc.idmapd daemon.
Debugging and Deploying Secure Linux NFS
Steve Dickson, *Red Hat*

Red Hat Summit 2006
Agenda

• Debugging tips
• Setting up Kerberos
Debugging Tips

• **System Oops or Panics**
  - Netdumps – dumps system core over a lan
    • Netdump and netdump-server
  - Diskdumps – dumps system cores to swap
    • Savecore, /etc/sysconfig/diskdump
  - Crash command for debugging live systems and system core dumps.
    • http://people.redhat.com/anderson/.crash_whitepaper/
  - Kernel-debuginfo RPMS need for crash.
    • http://people.redhat.com/duffy/debuginfo/index-js.html
Debugging Tips

• System or Application Hangs – Use System Request facility
  - Set kernel.sysrq=1 in /etc/sysctl.conf
  - On console, AltSysRq commands
    • AltSysRq-t – system wide backtrace
    • AltSysRq-m – dumps memory stats
    • AltSysRq-c – cause system core dump
  - /proc/sysrq-trigger
    • echo 't' > /proc/sysrq-trigger
Debugging Tips

• Application Failures on Live Systems
  - Ethereal network traces.
    • Use tethereal(1) instead of tcpdump(8)
    • Use -w to create binary capture file
    • Constrain what is being captured with 'host' argument
      Ex: tethereal -w /tmp/data.pcap host <nfsserver>
  - /var/log/messages
    • Error are generally logged
Setting up Kerberos Mounts

• Create machine credits on both the server and client
  
  - Use kadmin or kadmin.local to create an machine credit in /etc/krb5.keytab
    
    • addprinc -randkey nfs/pro5.redhat.com
    
    • ktadd -e des-cbc-crc:normal nfs/pro5.redhat.com
  
  Note: Only des-cbc-crc encryption type is supported.
Setting up Kerberos Mounts

- Create machine credits (continued)

  Use (as root) klist -k to verify the /etc/krb5.keytab is setup correctly.

```
pro5# klist -k
Keytab name: FILE:/etc/krb5.keytab
KVNO Principal

6 nfs/pro5.lab.boston.redhat.com@STEVED.COM
```
Setting up Kerberos Mounts

- Setup kerberos configuration file, /etc/krb5.conf
  
  - [realms] section
    
    STEVED.COM {
      kdc = kerberos.redhat.com:88
      admin_server = kerberos.redhat.com:749
    }
  
  - [domain_realm] section
    
    .steved.com = STEVED.COM
    steved.com = STEVED.COM
Setting up Kerberos Mounts

- Setup kerberos configuration file (continued)
  - In cross-realm environments client mappings must be set up in the [domain_realm] section.

  [domain_realm]
  pro5.redhat.com = STEVED.COM
  pro1.redhat.com = STEVED.COM
Setting up Kerberos Mounts

• Turn on SECURE_NFS
  - Added 'SECURE_NFS=yes' to /etc/sysconfig/nfs.
  - On the client, start rpc.gssd
    • service rpcgssd start
  - On the server start rpc.svcgssd
    • service rpcsvcsgssd start
  - Check /var/log/messages for start up errors.
    • To turn on debugging add -vvv to OPTIONS in start-up script
Setting up Kerberos Exports

- Use gss/krb5, gss/krb5i or gss/krb5p as the machine names in the export list.

```
/     *(ro,sync,fsid=0)
/home *(rw,sync,nohide,fsid=1)
/home gss/krb5(rw,sync,nohide,fsid=1)
/home gss/krb5i(rw,sync,nohide,fsid=1)
/home gss/krb5p(rw,sync,nohide,fsid=1)
```

- Either restart the NFS server or use 'export -r' to make kernel sees new exports.
Setting up Kerberos Mounts

• Common errors:
  
  − Sync up system clocks with NTP.
  − Use fully-qualified host names.
  − Make sure NFS is in the list of services in /etc/services

  nfs    2049/tcp    nfsd
  nfs    2049/udp    nfsd
References

• CITI NFSv4 Project – Univ of Michigan
  - http://www.citi.umich.edu/projects/nfsv4

• NFSv4 Test Matrix - OSDL
  - http://developer.osdl.org/dev/nfsv4
References (continued)

• The NFS version 4 Protocol
  - Presented at SANE 200. Written by Pawlowski, Shepler, Beame, Callaghan, Eisler, Noveck, Robinson and Thurlow.

• Linux NFS Version 4: Implementation and Administration
  - Presented at OLS 2001.
  - Written by William A Adamson (CITI) and Kendrick M. Smith
Open Discussion