NFS version 4 for FreeBSD

Chuck Lever, Network Appliance, Incorporated
Overview

- What is NFS version 4?
- After NFS version 4
- State of vendor NFS implementations
- Potential FreeBSD client road maps
What is NFS version 4?

- Improvements over NFS version 3
  - Real IETF specification
  - Internet friendly
    - Compatibility with firewalls
    - Real security
    - Fewer transactions on the wire
What is NFS version 4?

- Transport Implications
  - Reliable transports only
    - UDP is out
    - TCP and SCTP (datagram and stream) is in
  - Retransmit over stream transports specified
    - Usually no server-side DRC is needed
  - Futures
    - Sessions
    - Remote DMA
What is NFS version 4?

• RPC changes
  - Name mapping
    • No uids/gids on the network, only strings
  - Two NFSv4 RPC procedures supported:
    • NULL: Same as before
    • COMPOUND: Client can send complex and unique operations in a single round trip
  - Support for RPCSEC GSSAPI is now mandated by the NFS spec
    • Basic pseudoflavors are Kerberos and SPKM/lipkey
What is NFS version 4?

- No auxiliary protocols
  - Mountd, NLM, NSM, and NFS ACL all rolled into a single new protocol
  - Server uses single well-known port 2049
    - More firewall-friendly
What is NFS version 4?

• Stateful
  - Protocol now has OPEN and CLOSE
  - State and sequence IDs
  - State recovery required after reboots/panics

• Locking
  - Mandatory and advisory
  - Byte-range locking
  - Share reservation
  - Lease-based
What is NFS version 4?

- Delegation enables aggressive caching
  - Server allows a client to assume that there is limited or no sharing
  - File granularity (not mount point or byte)
  - Server calls back when other clients want access
  - Two types: read and write
  - Directories not yet supported
  - Similar to CIFS oplocks
What is NFS version 4?

- Other new features
  - ACLs
    - Richer than POSIX
    - Similar to, but not the same as, Windows NT
  - Named attributes
    - Multiple auxiliary unformatted byte streams
  - Server-side pseudofs
  - Migration & replication
    - Client/server protocol support
    - NOT server/server (could be NDMP, FTP, etc)
NFS Futures

- Standards work (NFS version 4.1)
  - Byte-range delegation
  - Directory delegation
  - NFS sessions
  - CCM
- Current research projects
  - Parallel NFS
  - 10GbE
  - RPC over Remote DMA
State of Implementations

- What's on the market today
  - Sun Solaris 10
  - IBM AIX 5.3
  - Red Hat's Enterprise Linux and Fedora Core
  - Network Appliance ONTAP 7.0
  - Hummingbird NFS version 4 for Windows
State of Implementations

• No implementation on the market is complete
  - Missing everywhere
    • Migration and replication
    • SPKM3
    • Named attributes
  - Missing in some implementations
    • Kerberos
    • NFS version 4 ACLs
Potential Development

- RPCSEC GSSAPI
  - Support for all versions of NFS
  - Easier to deploy Kerberos
  - Possible support SPKM/lipkey
Potential Development

• NFSv4
  - Integration of NFSv[23] and NFSv4 client
  - Support only basic NFSv4 features
    • No support for ACLs, delegation, migr/repl, *eg.*
  - Fully support NFSv4.0 spec
  - NFSv4.1 and beyond
Potential Development

- Not related to NFS version 4
  - IPv6
  - Automounter
  - Application API enhancements
    - direct I/O
    - efficient scatter/gather
    - async I/O
  - Improvements to NFSv[23]
    - Data and attribute caching semantics
    - NLM stability and performance
Potential Development

- Not related to NFS version 4
  - Performance improvements
    - Better cache coherency
    - Advanced transports
    - SMP
    - Observability tools
  - Maintainability
    - Bug fixing resources
    - Code reviewers
    - Testing infrastructure
Further Discussion