

NFS/RDMA Status

Tom Talpey

tmt@netapp.com

IETF-70 Vancouver

December 5, 2007

Draft status

- 3 drafts, all post- WG Last Call
- Now in IESG review
- Problem Statement
 - draft-ietf-nfsv4-rdma-problem-statement-07
- RPC/RDMA protocol
 - draft-ietf-nfsv4-rpcrdma-06
- NFS Direct mapping
 - draft-ietf-nfsv4-nfsdirect-06

Problem Statement

- Currently updating in response to IESG comments
- Clarifications, improvements
- However...
 - Security considerations need additional text
 - And the others need more
- Expect to modify Problem Statement and advance it separately (now)

RPC/RDMA and NFSDirect

- Some IESG comments pending
 - Overuse of acronyms
 - Numerous referenced docs now RFCs
 - Clarity, etc.
- Security considerations
 - Must make mandatory security requirement
 - Must map out how to implement
 - May (will) require new companion security draft

Security requirement

- Problem – how to maintain performance when RPCSEC_GSS integrity/privacy in force?
- Documents currently leverage RDDP security requirement for IPsec
 - RFC5042
- “Downgrade” to NULL wrap/unwrap when connection protected by IPsec
- Referenced CCM for this

Security requirement (2)

- CCM now published as RFC5056
 - Thanks, Sam H. and Nico!
- Still describes NFS/RDMA/RDDP layering
- No longer describes the NFSv4/RPCSEC_GSS downgrade
- Therefore, now need to bridge this gap

Gap-bridging options

1. Do what iSER/RDDP (RFC5046) does?
 - Maybe not, does not address RPCSEC_GSS
2. Define a new RPCSEC_GSS op?
 - RPCSEC_GSS_BIND_CHANNEL
 - See discussion on nfsv4 mailing list (Nico)
3. Define a new Channel Binding shim
 - To sit between RPCSEC_GSS and IPsec
 - Name the channel, encapsulate its properties
 - Resurrect old text from CCM draft?

TBD

- Approach is under discussion
- Security AD may be willing to sponsor individual draft
- Alternatively, may be Better-Than-Nothing-Security (BTNS) work item
 - At least, has dependencies on BTNS
- Watch this space.

NFS/RDMA prototyping

- Two full open source projects under way
- Linux
 - Client experimental in upcoming 2.6.24
 - Server under review for future 2.6.25 (TBD)
 - Linux OFA API layering
 - Full support of iWARP, Infiniband
- OpenSolaris
 - Client, server to appear “soon”
 - Infiniband only
- Yes, they interoperate!

NFS/RDMA and pNFS

- Demo'd at SuperComputing 07
 - Reno, NV November 12-16
- pNFS client/server from Linux open source project
- “spNFS” (simple pNFS) MDS daemon from NetApp
- NFSv3/RDMA Data Servers
- Excellent, scalable performance
 - 4 clients, 4 dataservers, 10GbE iWARP interconnect
 - ~600MB/s/link (client cached i/o)
 - 2.4GB/s aggregate bw – linear scaling!

Watch this space

- For Linux NFS/RDMA client and server
- For OpenSolaris NFS/RDMA client and server
- For open source spNFS daemon
- For continued joint performance leverage between pNFS and NFS/RDMA

Backup

- This overview was assembled and brought to you by...

NFS/RDMA recommends... ;-)



December 5, 2007

IETF70 Vancouver

13