NFSv4.1: An update

Mike Eisler
Senior Technical Director
2009-02-23
Connectathon 2009

mre dash ietf at eisler dot com

NetApp™

Go further, faster™
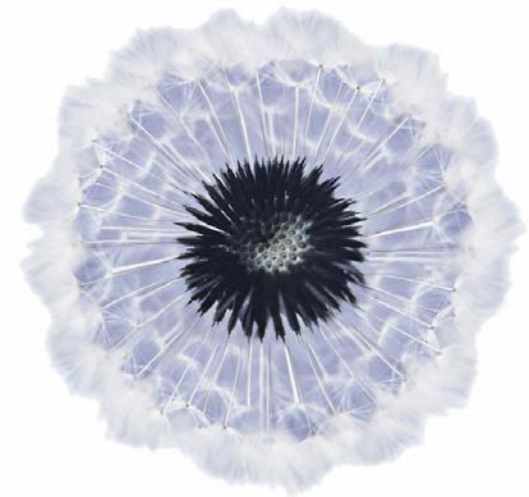
Topics

- Briefly about the NFSv4 working group
- Status of current work items
  - including NFSv4.1
- After NFSv4.1
- Non-Topic:
  - A detailed description of NFSv4.1
    - Go to previous Connectathons for presentations
      - Eisler
      - Noveck
      - Shepler

Briefly about the NFSv4 working group (WG)

- Stimulus was Sun giving change control of the NFS protocol to IETF
  - http://www.rfc-editor.org/rfc/rfc2339.txt
- WG was chartered in 1998 to specify NFSv4
  - RFC3010 published in 2000
    - DOA specification for NFSv4.0
  - RFC3050 published in 2003
    - All life signs are normal
- Since then WG has picked up:
  - ONC RPC bis items
    - ONC RPC WG was concluded in 2001
    - bis: Latin word meaning "twice"
      - bis documents are revisions necessary to allow a document to advance from Proposed Std. to Draft Std, or Draft Std. to Standard
    - RDMA
    - NFSv4.1
    - Federated File System

NFSv4.1: differences from NFSv4.0

- Introduces a session layer
  - Exactly once semantics
    - vastly simplifies locking
  - Trunking
    - utilize more paths
    - retry failed requests over other paths
- Brings delegations to directories and symlinks
- OPEN enhanced to specify ACL on created files
- Delegations are easier to manage
  - recall any semantics
 - re-acquisition without re-open
- pNFS

pNFS positioning: one editor's opinion

- Performance acceleration is not the goal
- pNFS (especially when combined with striping) simplifies storage and data management
  - data is sprayed evenly across storage cluster
    - load and capacity balancing occur without customer intervention
  - pNFS layout recall allows data to be re-distributed across cluster
    - re-size cluster
    - re-tier cluster

Better Management Through pNFS

No Striping

Striping

No pNFS

With pNFS

Clients

Status of Current Work Items
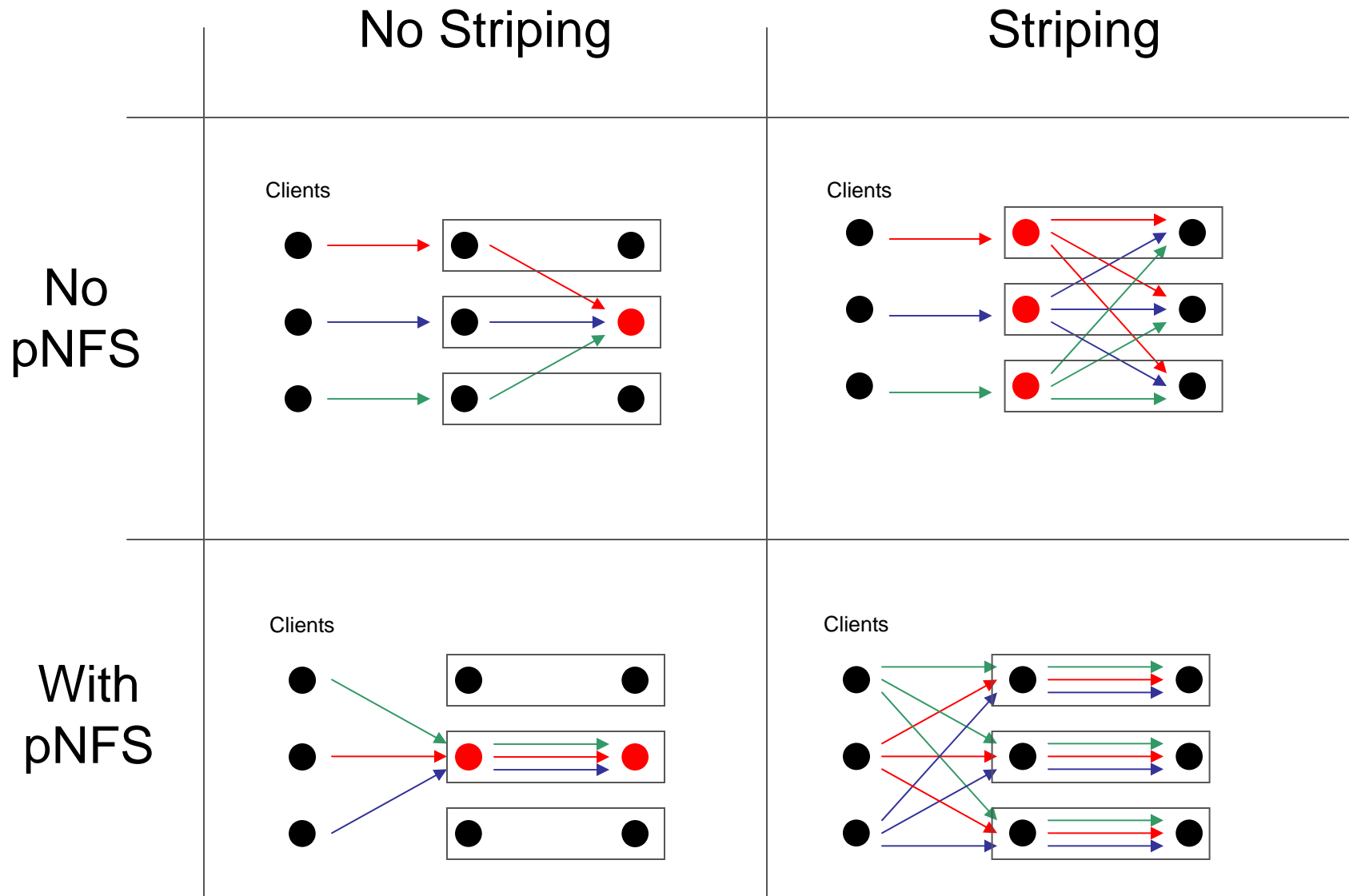
- http://tools.ietf.org/wg/nfsv4 is your friend
- Work Items
  - Active:
    - draft-ietf-nfsv4-federated-fs-admin-* Active
  - Approved:
    - draft-ietf-nfsv4-rfc1831bis – Approved ‡
    - This moves ONC RPCv2 to Draft Standard
  - In RFC Editor's Queue:
    - draft-ietf-nfsv4-minorversion1-dot-x *
    - draft-ietf-nfsv4-minorversion1 *
    - draft-ietf-nfsv4-nfs-rdma-problem-statement †
    - draft-ietf-nfsv4-nfsdirect †
    - draft-ietf-nfsv4-pnfs-block *
    - draft-ietf-nfsv4-pnfs-obj *
    - draft-ietf-nfsv4-rpc-netid ‡ §
    - draft-ietf-nfsv4-rpcrdma †
  - Published:
    - draft-ietf-nfsv4-rpcsec-gss-v2 §
      - http://tools.ietf.org/html/rfc5403

* NFSv4.1/pNFS documents
† NFS/RDMA documents
‡ Documents required to publish NFSv4.1, pNFS documents
§ Documents required to publish RDMA documents

Life in the Editor's Queue

- The document author provides the RFC Editor XML source
- The RFC Editor eventually sends edited xml and preliminary RFC
  - Editor/Author has an indefinite amount of time to review, rebut
- NFSv4.1 is the longest document in IETF history
  - Expect it to be the longest wait for an RFC in IETF history
  - You can implement from http://tools.ietf.org/wg/nfsv4/draft-ietf-nfsv4-minorversion1/
  - Lots of opportunity to find errors and immortalize your name in the RFC

After NFSv4.1

- Work Items:
  - Federated File System
    - http://www.ietf.org/proceedings/08nov/slides/nfsv4-1.pdf
- Currently Proposed Work Items
  - MAC Labeling
    - http://www.ietf.org/proceedings/08nov/slides/nfsv4-2.pdf
    - draft-quigley-nfsv4-sec-label
  - Metadata Striping
    - http://www.ietf.org/proceedings/08nov/slides/nfsv4-3.pdf
    - draft-eisler-nfsv4-pnfs-metastripe
  - De-Duplication
    - http://www.ietf.org/proceedings/08nov/slides/nfsv4-3.pdf
    - draft-eisler-nfsv4-pnfs-dedupe
- All the above either require zero to little change to NFSv4.1
- MAC Labeling will likely require RPCSEC_GSSv3

Thanks

Q/A