



Go further, faster™

# Extending NFS to Support Enterprise Applications

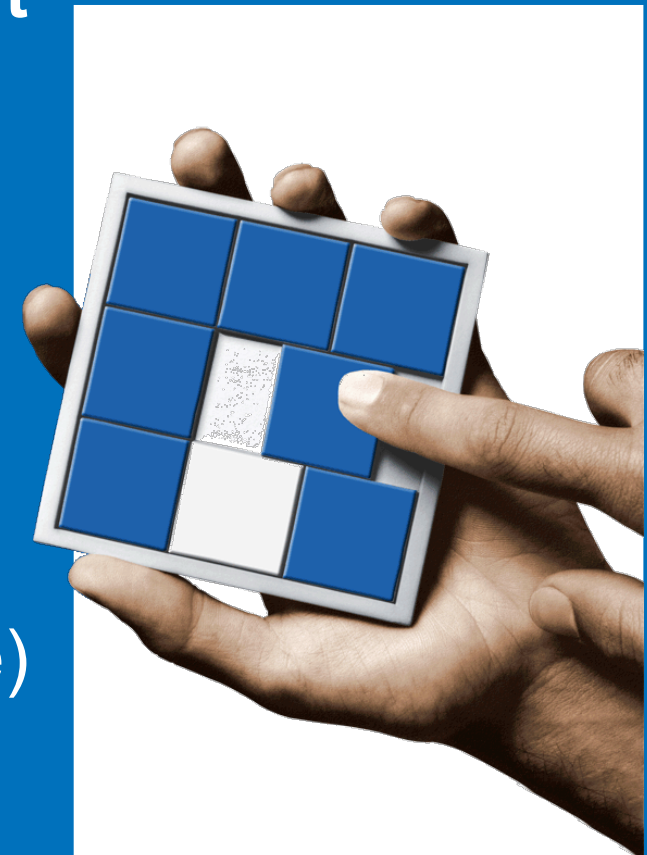
draft-eisler-nfsv4-enterprise-apps-01

IETF-79

2010-11-09

Mike Eisler (NetApp)

Margaret Susairaj (Oracle)





# draft-eisler-nfsv4-enterprise-apps-01

- Differences from last update: N/A
- NFSv4.x or standalone?
  - Mostly NFSv4.x; possibly some of this is errata for RFC5661
- Key points solved
- Issues of contention
- Data that supports proposal
- Overlaps
- Should proposal be combined/not combined?
- Who should review?
- Next steps



# draft-eisler-nfsv4-enterprise-apps-01

- Key points solved
  - Addressing needs of applications like databases which have higher requirements for data integrity, deterministic QoS, and availability
  - INITIALIZE operation
    - A hole punch with a specific pattern
      - because patterns of zeroes are among the most common forms of data corruption
  - IO\_ADVICE operation
    - Provides advice to server regarding intended I/O pattern to file
  - READ\_WITH\_ADVICE/WRITE\_WITH\_ADVICE
    - For those cases where a COMPOUND combining IO\_ADVICE and READ or WRITE do not suffice. E.g. an IO\_ADVICE to not cache data returned by READ.
    - Anticipates 100Gbit/sec links and defines 64 bit byte counts
    - Adds a LAYOUT\_SYNC4 stable\_how indication in WRITE\_WITH\_ADVICE response to indicate that a LAYOUTCOMMIT is not necessary



# draft-eisler-nfsv4-enterprise-apps-01

- Key points solved
  - SET\_WORKFLOW\_TAG
    - Sets a workflow tag on a session
    - Opaque to protocol, interpreted by client and server
      - One possible interpretation is as a priority
  - SESSION\_CTL
    - Allows client to request change in session parameters
    - Primary use case is size of slot table on fore channel
      - NFSv4.1 slot table introduced a limit on number of outstanding operations
      - This limit might be much less than the number of operations an application has in progress



# draft-eisler-nfsv4-enterprise-apps-01

- Key points solved
  - New EXCHGID4\_FLAG\_SUPP\_FENCE\_OPS capability
  - NFSv4.1 specification leaves potential for dangling operations when a session is destroyed or an EXCHANGE\_ID destroys all state of a previous instance of a client-owner
  - The new capability makes it clear that the server will not allow dangling operations



# draft-eisler-nfsv4-enterprise-apps-01

- Issues of contention
  - Is the EXCHGID4\_FLAG\_SUPP\_FENCE\_OPS capability necessary or is this just RFC5661 errata
    - Unfortunately, few NFSv4.1 servers deal with dangling ops of killed sessions and client IDs
  - Is support for 128 bit file counts over kill?
  - Should we extend XDR to add a 128 bit integer?
- Data that supports proposal
  - The experience of database vendors and storage vendors around these issues, especially with block protocols
    - The answer: “so go install a SAN or iSCSI” takes protocol choice from customers
      - Many customers have chosen NFS for these workloads
- Overlaps
  - HOLEPUNCH operation from draft-iyer-nfsv4-space-reservation-ops-01.txt is a subset of INITIALIZE

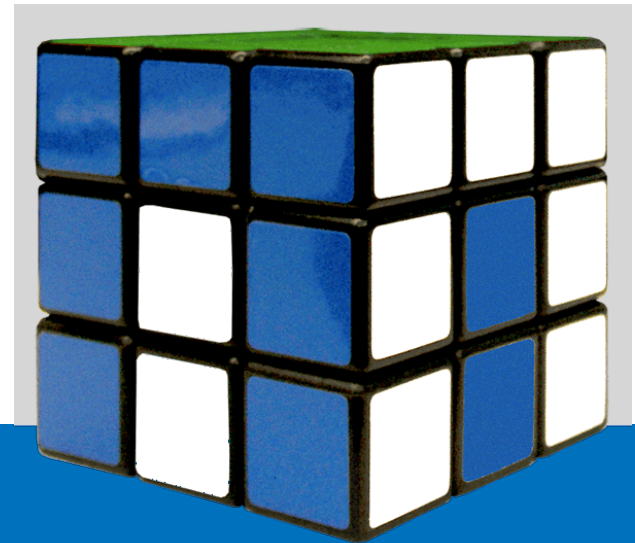


## draft-eisler-nfsv4-enterprise-apps-01

- Should proposal be combined/not combined?
  - Should be part of draft-ietf-nfsv4-minorversion2
- Who should review?
  - WG
- Next steps
  - Oracle and NetApp are likely to prototype this

# Thanks

draft-eisler-nfsv4-enterprise-apps-01



## Questions and Answers