NFSv4.1 Sessions: An MPTCP Use Case

Chuck Lever <chuck.lever@oracle.com>
Andy Adamson <andros@netapp.com>

NFSv4.1 Sessions

- A bounded server-side RPC replay cache, enabling Exactly-Once RPC execution semantics (RFC 5661 Section 2.10)

- One or more sessions are created by each NFSv4.1 client

- One or more transport connections are bound to each session

- Sessions outlive transport connections

Trunking

- RFC 5661, Section 2.10.5

- Multiple transport connections from one client

  - High availability (*e.g.*, path failover)

  - Performance (*e.g.*, parallel connections and NICs)

  - Hybrid transports (*e.g.*, TCP and RDMA)

- No native facility in NFS for server interface discovery

Forms Of Trunking

- pNFS

- Client trunking

  - Client creates multiple sessions

  - One per mount; or one for I/O, one for metadata

- Session trunking

  - One session, multiple transports

Trunking With MPTCP

- A single MPTCP socket bound to a session provides similar benefits, but:

  - Only TCP transports are supported

  - RDMA not included; uses verbs, not sockets

  - No information about characteristics of individual flows

  - No control over traffic routing among subflows