



SDC¹⁸

September 24-27, 2018
Santa Clara, CA

www.storagedeveloper.org

SMB3-NFS Gateway over Lustre

Ujwal Bhambani

Vinit Agnihotri

Ryussi Technologies

SMB3-NFS Gateway over Lustre

- ❑ Motivation for SMB3-NFS Gateway on Lustre
- ❑ Architectural Goals
- ❑ SMB3-NFS over Lustre Architecture
- ❑ Unified SMB3 & NFSv3/4 stacks – Challenges
- ❑ Choice of NFS server
- ❑ Choice of SMB server
- ❑ Unified SMB3 & NFSv3/4 stacks – Solutions
- ❑ Key Takeaway

Motivation for SMB3-NFS Gateway on Lustre

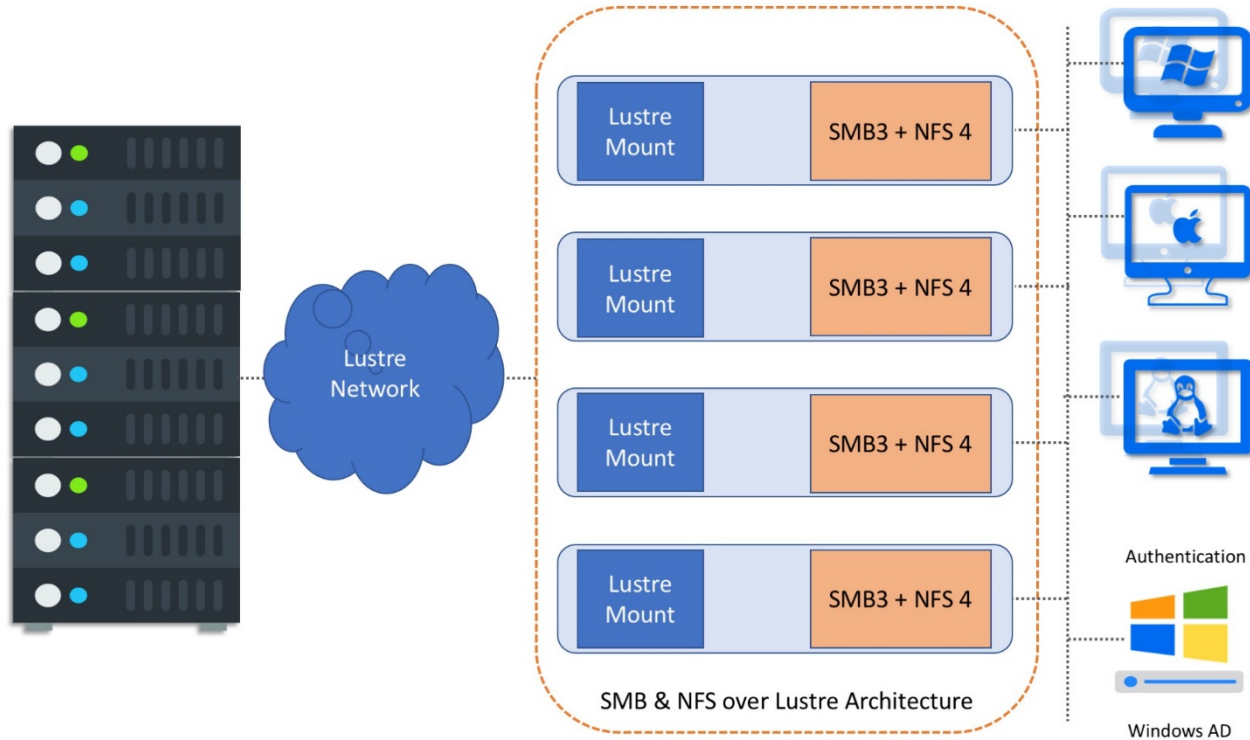
- ❑ Lustre storage available to non-Lustre clients
- ❑ Data Visualization – clients Windows, Mac, Linux
- ❑ Data Acquisition – non-Lustre clients



Architectural Goals

- ❑ Security
- ❑ High Availability
- ❑ Manageability
- ❑ Performance

SMB3-NFS GW over Lustre architecture



Unified SMB-NFS GW Challenges

- ❑ Security
 - ❑ Authentication
 - ❑ Authorization
- ❑ File Sharing mode
- ❑ File Locking (byte-range)
- ❑ Common Metadata & data cache across protocols
- ❑ Unified Configuration & Management

Advanced Feature

- ❑ SMB Leasing(File & Directory) or Delegation in NFS v4
 - ❑ Mechanisms to Break lease/delegation
 - ❑ SMB/NFS timeouts

Unified SMB-NFS GW Challenges

- ❑ Cluster configuration (AP, AA)
- ❑ Unified metadata cache across nodes
- ❑ Mechanism to notify file/directory changes across nodes
- ❑ Unified Installation
- ❑ Unified Configuration & Management across nodes

Authentication and authorization - challenges

- ❑ Methods available
- ❑ Mix environment
- ❑ User mapping

Unified SMB-NFS Gateway - a potential solution & limitations

NFS & SMB acquire linux file-system level locks

- ❑ File Sharing
- ❑ File Locking

Only advisory lock support possible

File sharing modes – challenges of using linux file-system level locks

SMB file sharing modes

- ❑ File share read, write, delete

NFS file sharing modes

- ❑ `Open4_share_access_read,write,both`

File Locking modes – challenges of using linux file-system level locks

File Locking mode challenges

- ❑ Linux file system call `fcntl()`
- ❑ Linux fd based lock `flock()`
- ❑ Non-posix locks (Open file description locks)
 - ❑ `f_ofd_setlk, getlk`

Unified configuration - challenges

- ❑ Necessity of unified configuration
- ❑ Config options
- ❑ Utilities
- ❑ Installation

Which NFS server to be integrated?

NFS kernel Or NFS Ganesha

Criteria for selection could be the following features

- ❑ NFS versions supported
- ❑ Easy maintenance of code
- ❑ Modularity
- ❑ Interoperability with other protocols
- ❑ Performance

NFS-Ganesha

NFS Ganesha

- ❑ User mode
- ❑ Maintainability
- ❑ Modularity
- ❑ Easy to configure

Which SMB server to be integrated?

Any SMB server that supports (samba, MoSMB,...)

- ❑ SMB 2.* & 3.* protocols
- ❑ Persistent Handles v1 & v2
- ❑ Transparent failover of SMB connections
- ❑ File locking
- ❑ File sharing modes
- ❑ Interoperability with other protocols
- ❑ Performance
- ❑ SOFS support

Unified SMB-NFS Gateway - I

A potential Authentication solution

- ❑ Authentication
 - ❑ Kerberos based authentication for SMB and NFS
- ❑ Authorization
 - ❑ NFS supports POSIX ACLs
 - ❑ SMB to translate NT ACLs<->POSIX ACLs
 - ❑ SMB to support AD users mapping to Linux uid/gid.

SMB/NFS using Linux file-system level locks

– limitations

- ❑ flock(), fcntl() calls in linux
- ❑ Locking calls advisory in nature
- ❑ SMB semantics cannot be implemented
- ❑ Mandatory locking not available
- ❑ Unable to provide Cluster-wide support

Unified SMB-NFS Gateway – a potential solution

Distributed middleware

- ❑ Global view of data structure
- ❑ Mandatory Locking support can be provided
- ❑ Every open/create (SMB/NFS) shall result in mandatory locking
- ❑ Common locking across cluster is possible
- ❑ NFS grace period could be honored by SMB, to avoid possible lock stealing.

Unified SMB-NFS Gateway – a potential solution

Distributed middleware

- ❑ SMB Directory & File Leases and NFS Delegation could co-exists
- ❑ Lease break and delegation recalls across cluster can be possible
- ❑ Lease break and delegation recalls across SMB and NFS

Unified SMB-NFS – a potential solution

- ❑ Active-Passive
 - ❑ Network Load Balancer VIP (virtual IP)
- ❑ Active-Active cluster nodes
 - ❑ Network Load Balancer VIP (virtual IP) with load-balancer
 - ❑ DNS RR
- ❑ Unified Installation
- ❑ Unified Configuration & Management

Key Takeaways

- ❑ Some earlier solutions to this problem suggest modifications to underlying file-system to accommodate the challenging scenario of cross-protocol support.
- ❑ We propose to solve this using a common distributed middleware facilitating the interoperability between protocols as a file-system independent approach.
- ❑ This approach is applicable as well to object store exported over NAS 😊

References

- ❑ Lustre
 - ❑ <http://lustre.org/>
- ❑ NFS
 - ❑ <https://tools.ietf.org/html/rfc3010>
 - ❑ <https://nfs-ganesha.github.io/>
- ❑ SMB
 - ❑ <https://www.samba.org/>
 - ❑ <https://www.mosmb.com>

All product and company names are trademarks™ or registered® trademarks of their respective holders. Use of them does not imply any affiliation with or endorsement by them.