



SDC 18

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 (standalone/cluster mode) – Challenges
 - ❑ Choice of NFS server
 - ❑ Choice of SMB server
 - ❑ Unified SMB3 & NFSv3/4 stacks (standalone/cluster mode) – Solutions
- Breadth first, not depth approach

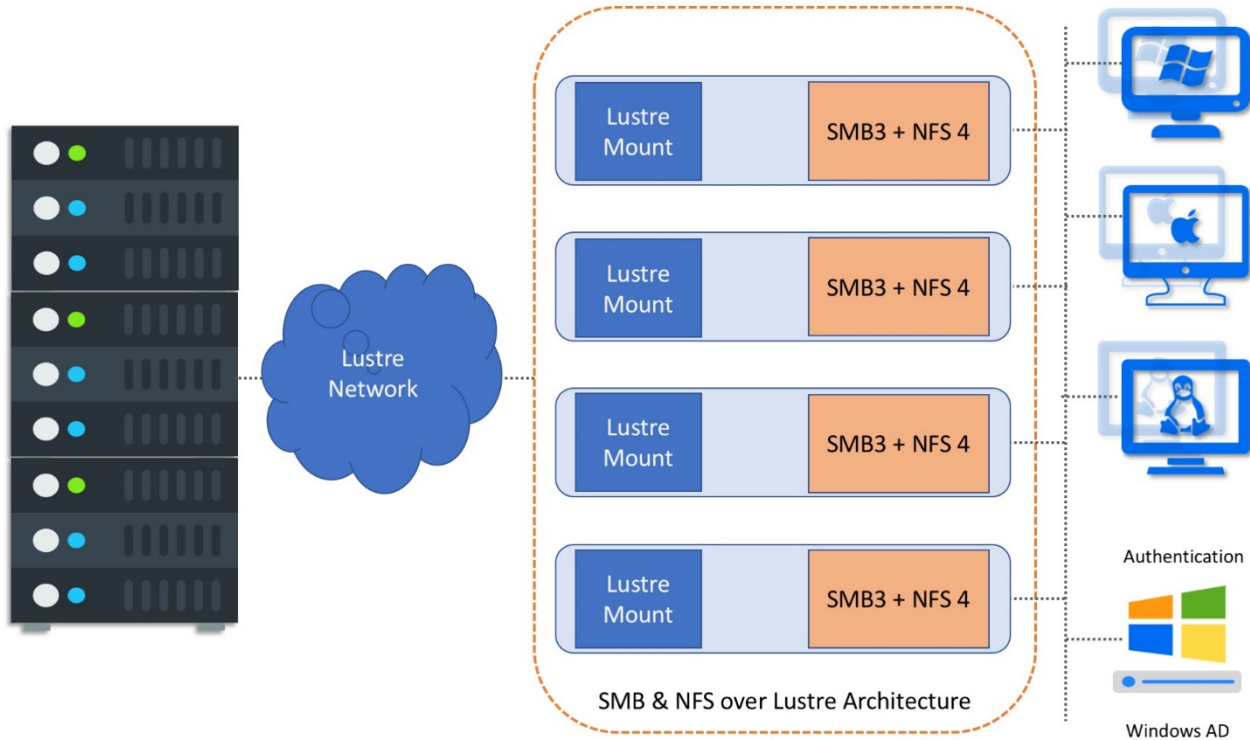
Motivation for SMB3-NFS Gateway on Lustre

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

Architectural Goals

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

SMB3-NFS over Lustre architecture



Unified SMB-NFS Challenges (Standalone Server)

- ❑ User Security
 - ❑ Authentication
 - ❑ Authorization
- ❑ File Sharing mode
- ❑ File Locking (byte-range)
- ❑ Metadata & data cache, if used by any stack must be unified across protocols, or disabled
- ❑ 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 Challenges (Cluster)

- ❑ Cluster configuration (AP, AA)
- ❑ Unified metadata cache across nodes
- ❑ Unified Installation
- ❑ Unified Configuration & Management across nodes

Which NFS server to be integrated?

NFS kernel vs NFS Ganesha

Criteria for selection could be the following features

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

Which NFS server to be integrated? (contd.)

NFS Ganesha

Pros:

- ❑ NFS versions supported v3, v4
- ❑ User mode code, Easy maintenance of code
- ❑ Modularity

Cons:

- ❑ Performance degrades on disabling the inode & data cache

NFS Kernel

Pros:

- ❑ NFS versions supported v3, v4
- ❑ Performance

Cons:

- ❑ Easy maintenance of code
- ❑ Modularity

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

Unified SMB-NFS Gateway (Standalone) – I a potential solution

- ❑ Authentication
 - ❑ SMB supports AD kerberos authentication
 - ❑ NFS supports kerberos authentication
- ❑ Authorization
 - ❑ NFS supports POSIX ACLs
 - ❑ SMB must translate NT ACLs<->POSIX ACLs
 - ❑ AD users mapped to Linux uid/gid must be supported by SMB server stack. File access checks must be performed accordingly

Unified SMB-NFS Gateway(standalone) – II a potential solution

- ❑ File Sharing & File Locking
(Advisory lock support across SMB & NFS)
 - ❑ NFS must acquire file-system level lock
 - ❑ SMB must acquire file-system level lock
- ❑ Unified Configuration & Management
 - ❑ CLI based
 - ❑ UI based

Unified SMB-NFS Gateway(standalone) – III a potential solution

- ❑ Mandatory SMB Locking support
 - ❑ SMB & NFS must have common distributed middleware
 - ❑ Checks need to be made before each file open
- ❑ SMB Directory & File Leases and NFS Delegation
 - ❑ SMB & NFS must have unified distributed middleware
 - ❑ Checks need to be made before each file open

Unified SMB-NFS Challenges (Cluster)

- ❑ Active-Passive
 - ❑ Network Load Balancer VIP (virtual IP)
- ❑ Active-Active cluster nodes
 - ❑ Network Load Balancer VIP (virtual IP) with load-balancer
 - ❑ DNS RR
- ❑ SMB & NFS must have common distributed middleware across nodes for locks & leases
- ❑ Unified Installation
- ❑ Unified Configuration & Management

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.