Building a Highly Scalable and Performant SMB Protocol Server

Dr. Sunu Engineer
CTO, Ryussi Technologies
About Ryussi

• Creator of MoSMB – SMB with Mojo.
• Specializes in Storage and Systems Software Product Development and Consulting.
• Founded in 2012 by Storage and Networking veterans.
• Has solved several challenging problems in the areas of Storage, Virtualization, OpenStack, SDN and Big Data Analytics.
• Clientele include Fortune 500 companies and several startups.
About MoSMB

• Lightweight, ANSI-C SMB 2/3 stack on Linux variants built completely from ground up.
• Architected to support high performance, high scalability and continuous availability.
• Enterprise class feature set to support common SMB use cases such as Microsoft Hyper-V.
• Flexible architecture with custom interfaces to integrate into diverse storage stacks quickly and efficiently.
• Complete ecosystem support including SMB Direct, ODX, RVSS, clustering and witness protocol, SQOS, SMI-S management, etc.
Focus & Content of this presentation

- SMB server Design & Architecture
- Scalability – design considerations
- Performance – design considerations
- Comparative Performance numbers
- SMB server & Cluster support

This presentation is about MoSMB – SMB with Mojo
MoSMB Design Goals

<table>
<thead>
<tr>
<th>Category</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Performance</td>
<td>Lightweight, Best in class performance features (RDMA, Multichannel, Directory Leasing)</td>
</tr>
<tr>
<td>High Scalability</td>
<td>Scale-Out Architecture, Capable of taking advantage of multi processor, multi core systems</td>
</tr>
<tr>
<td>Transparent Failover/Continuous Availability</td>
<td>Witness Protocol Support, Persistent Handles</td>
</tr>
<tr>
<td>Enterprise Level Security</td>
<td>End-to-end Encryption, Secure Dialect Negotiation, AES-CMAC Signing</td>
</tr>
<tr>
<td>Ease of integration</td>
<td>ANSI C, Modular, Pluggable Architecture, Well Defined Interface</td>
</tr>
<tr>
<td>Full SMB Ecosystem</td>
<td>SMB Direct, ODX, RVSS, Witness, Clustering, SQOS, SMI-S</td>
</tr>
</tbody>
</table>
Scalability Design Considerations

- A single light-weight process
- Event driven architecture
- Thread pool model
- Minimal Resource Usage
- Zero-copy model
- A well bounded utilization of resources on the server side irrespective of the load.
SMB Server Model

Client sends SMB request to MoSMB

Event loop is woken up by OS, queues request

SMB responses are sent to client

Threads processing SMB packets

Event loop returns result to client

SMB packets
Performance Design Considerations

- **Lean Architecture**
  - Low memory footprint
  - Event driven architecture
  - Zero copy model

- **Multi-channel**
  - Bandwidth aggregation using multiple NICs
  - Receive-side scaling support

- **SMB Direct (RDMA)**
  - Scalable, fast and efficient storage access
  - High throughput with low latency
  - Minimal CPU utilization for I/O processing
  - Load balancing, automatic failover and bandwidth aggregation via SMB Multichannel

- **Directory Leasing**
  - Caching file and directory metadata
Comparison of MoSMB vs Samba Performance

Server details: MoSMB & samba
Hw used: Intel(R) Xeon(R) CPU E3-1220 V2 @ 3.10GHz
Quad core, 8MB RAM
OS: Ubuntu 14.04
Client details: Windows 2012R2
Comparison of MoSMB vs Samba Performance

Blocksize vs Throughput(KB/s) for random read 50% and write 50%

Server details: MoSMB & samba
Hw used: Intel(R) Xeon(R) CPU E3-1220 V2 @ 3.10GHz
Quad core, 8MB RAM
OS: Ubuntu 14.04
Client details: Windows 2012R2
Comparison of MoSMB vs Samba Resource Utilization

CPU usage SAMBA vs MoSMB

System memory usage SAMBA vs MoSMB

Server details: MoSMB & samba
Hw used: Intel(R) Xeon(R) CPU E3-1220 V2 @ 3.10GHz
Quad core, 8MB RAM
OS: Ubuntu 14.04
Client details: Windows 2012R2
FIO Job Description: Filesize=500m ,IODEpth=8,Number of jobs=1,Number of clients=1,Blocksize=64k
SMB server & Cluster support

- MoSMB supports a well defined cluster interface
- Can integrate with varied underlying cluster software
- Uses an internal RAFT based implementation
- Witness Server for Continuous Availability & Transparent Failover
SMB Scale-out

- Active/Active file shares
- Fast failure recovery
- Zero downtime
- Witness Server support
- DFS-R support
Thank You

Ryussi Technologies Pvt. Ltd.,
www.ryussi.com
www.mosmb.com
info@ryussi.com