MOSMB SCALABILITY ARCHITECTURE

Futuristic SMB3 stack for forward thinking companies

MOSMB (SMB WITH MOJO) IS RYUSSI'S ADVANCED SERVER MESSAGE BLOCK (SMB) SERVER COMPONENT LIBRARY OFFERING FOR THE DATA STORAGE MARKET TO ENABLE FILE SHARING IN A HETEROGENEOUS ENVIRONMENT THAT INCLUDES BOTH WINDOWS AND NON-WINDOWS COMPUTERS.

A key design goal of MoSMB has been to create a highly scalable server. This translates into using a minimal set of resources to provide that maximum amount of low latency, high throughput file based access to largest number of clients. With the modern architecture principles of creating distributed and scalable servers to MoSMB, some of the key design approaches are enumerated below:

A SINGLE LIGHT-WEIGHT PROCESS

MoSMB runs as a single process and uses threads for parallelization. This implies that there is very less overhead in establishing, terminating and switching between threads.

EVENT DRIVEN ARCHITECTURE

MoSMB uses an asynchronous, non-blocking approach which allows much better scalability since the I/O operations can overlap better with other I/O and non-I/O operations.

MINIMAL RESOURCE USAGE

MoSMB has a very low memory footprint per connection, per session, per open file handle data structure including the oplocks and leases and offers a well bounded utilization of resources on the server side irrespective of the load.

CONFIGURABLE THREAD POOL

MoSMB uses a configurable thread pool to restrict the overhead of frequent creation and destruction of threads.

SCALING PARAMETERS AND LIMITS

The parameters which affect the structural scaling of MoSMB as well (as the performance scaling) are mentioned below.

- Number of SMB clients

- Number of files open

- Number of concurrent connections

- Number of concurrent sessions

- Number of shares

- Number of MoSMB servers supported in a Scale-out cluster

The limits in the MoSMB framework are expected to be influenced significantly due to external subsystems – transport, storage, system memory, CPU, etc. The scalability numbers for MoSMB on an X86 quad core CPU with 32 GB are given in the table below:

Scalability Test	Tested in Ryussi-Lab	Software Limit	Comments
No of concurrent SMB client connections	8,000	None	Scales with CPU, RAM
No of concurrent sessions	8,000	None	Scales with CPU, RAM
No of SMB shares	5,000	None	Scales with CPU, RAM
No of connections accessing a share	8,000	None	Scales with CPU, RAM
No of files open	1,000,000	2^{64}	Scales with CPU, RAM
No of cluster nodes	32	1023 (configurable)	Depends on nodes available

CLUSTERED STORAGE ARRAY VIEW

The storage arrays form a clustered storage which exposes a uniformly accessible space. The set of

MoSMB servers form an active-active scale-out cluster that deal with same set of shares exported from the storage array. MoSMB has lockless shared database to support persistent handles v1 & v2 feature.
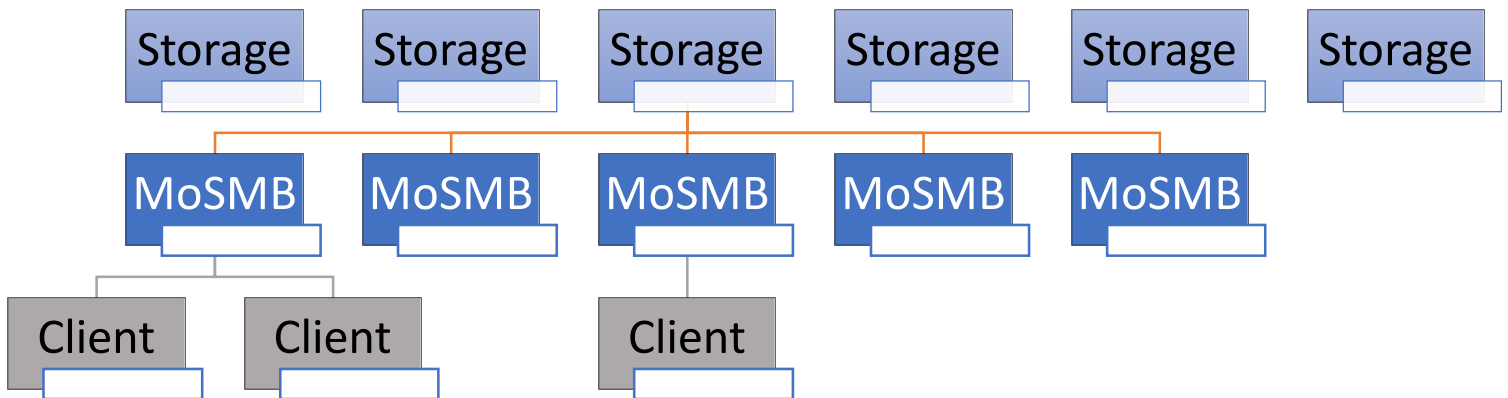
In order to produce high throughput and io parallelism from the system as a whole, one creates a set of servers who collectively serve out all the shares on the underlying storage. The underlying storage can be a clustered storage – clustered file system, object store like Ceph, fabric block storage like NVMf. The clients connect to different MoSMB servers in the cluster that effectively provide an aggregate throughput.

Witness server can be used to provide application level transparent failover in a clustered scenario.

To know how well MoSMB will scale on your platform, please contact sales@ryussi.com