How Scalable is your SMB?

Mark Rabinovich
Visuality Systems Ltd.

What is this all about?

- Visuality Systems Ltd. provides SMB solutions from 1998.
- NQE (Embedded) is an implementation of SMB client/server for the embedded world:
  - Consumer devices: printers, MFP, routers, smart devices, etc.
  - Industrial Automation, Medical, Aerospace and Defense
  - Anything else that is neither PC, MAC or Samba.
- NQ Storage is an SMB server implementation for Storage platforms.

This presentation is to share our experience with achieving the best scalability. All measures were done with NQ Storage not NQE.

Presentation Plan

- What is scalability?
- Load-wise scalability
- Usage-wise scalability
- Platform-wise scalability
- Measuring scalability
- Conclusions

What is scalability?

What a formal definition says?

The Oxford Dictionaries:

"*…the ability of…to be used…in a range of capabilities"*

The keyword is "range"

Wikipedia:

"*…is the capability of…to handle a growing amount of work"*

The keyword is "growing"

Dictionary.com: "*…the ability of something…to adapt to increased demands"*

The keyword is "increased"

… We can comprehend "scalability" from different points of view. Even a generic definition may emphasize a particular aspect of "scalability".

A Buzzword or a Value?

- The word “scalability” is used too much nowadays without a proper understanding of its meaning
- It is definitely a buzzword
- …but it also has a value

We will try discovering this value when it comes to SMB.

Scalability of an SMB Server

- To serve a range of usages
- To handle a range of loads
- To run on a range of platforms == portability

The three approaches are applying the same requirements but from different points of view.

Performance or Scalability?

- Customers are looking for both scalability and performance. Which is the primary one?
- Are they co-related? The answer is *YES*.
- Scalability is the performance exposure.
- There is no *real* performance without scalability.
  - “Scalability is the crust of performance…” says Sam Widerman, the CEO of Visuality Systems LTD.
- Lets talk *Scalability*.

Usage-wise scalability

Usage-wise or Market-wise?

Class	Examples
*) Embedded	Router Smart device Medical equipment
SOHO	Small office School
Dedicated	FAX server Document server
Media	Media editing studio Streaming
Storage	Data Center Cloud

*) Out of this session's scope (we do not need to be THAT scalable)

Customer case A

- The product is Virtual Data Center.
- Usage – mid-range Data Centers, Clouds.
- Simultaneous clients – 100 at least.
- Performance – a mix of 10,000 IPS per a core, performance should scale linearly when adding more cores.

Customer case B

- The product is High Performance Storage.
- Usage – top-level Data Centers
- Simultaneous clients – 10,000 at least.
- Performance – 30,000 IPS per a core, performance should scale linearly when adding more cores.

Customer case C

- The product is Midlevel Storage.
- Simultaneous clients – 1,000 at least.
- Performance – “as required by the market …”

Customer case D

- The product is a basic NAS.
- Simultaneous clients – 50 at least.
- Performance – “the best one can have out of a 1Gbs/10Gbs connection…”

Load-wise scalability

Load classification

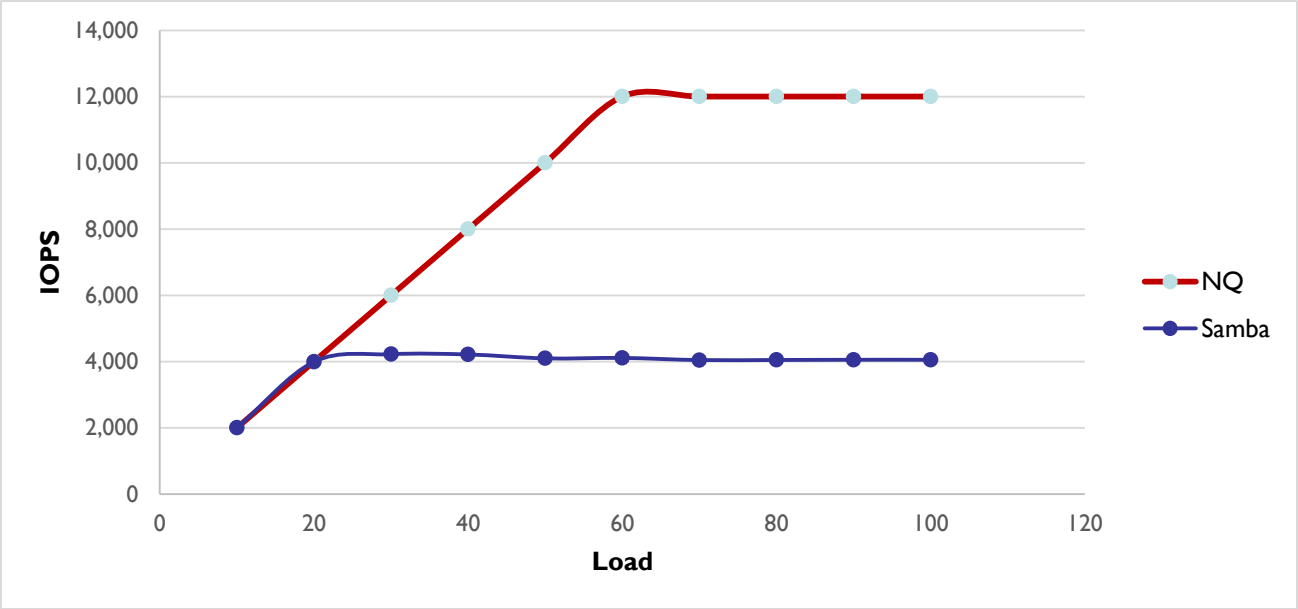
Class	Typical usage	Typical client application	Users	Traffic
Embedded	Green FAX	FAX Viewer	Few users or even single user	Low Various packet sizes
SOHO	Home NAS	MS Word	Several	Average Various packet sizes
Media editing	Media server	Media Editor	Few	Intensive Huge chunks
Cloud	File hosting	Drop box	Up to 10,000	Intensive Huge chunks
Storage	Data Center	HyperV, SQL Server	Up to 100,000	Intensive, ~4KB chunks

The same SMB Server solution may be used on each of the above markets (except from, maybe, Embedded) providing that SMB Server can be scaled to each of them.

The expectations

- With load increase, SMB Server is expected to provide at least the same value of IOPS.
- We expect IOPS increase in the low load area and saturation in the top (the next slide).
- On a very high load we will apparently face degradation.
- On even higher load we may expect denial of service.

The results (moderate load)



- Measured with SPEC SFS.2014
- Load = number of concurrent client applications

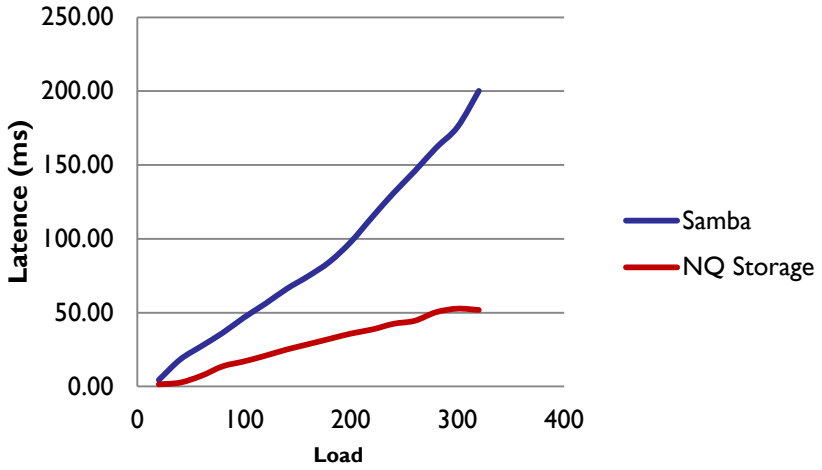
- Two different implementations of SMB are showing a similar trend.
- A later saturation point means less resource consumption.

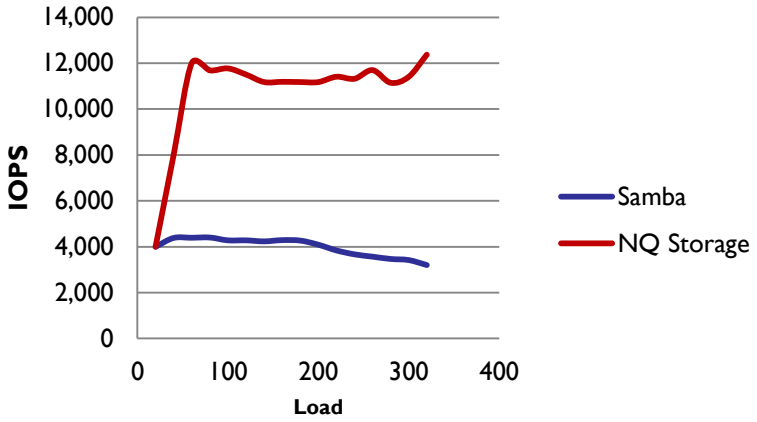
What happens on a very big load?

What is “Very high load”?

Load type	Notes	Expectations	Results
Full consumption of the bandwidth	10/40/50/100Gbs	Degradation of IOPS	10Gbs- see below
Numerous client connections	10,000 to 80,000 connections	Saturation of IOPS Connection failures	10Gbs - see below

The results (high load)



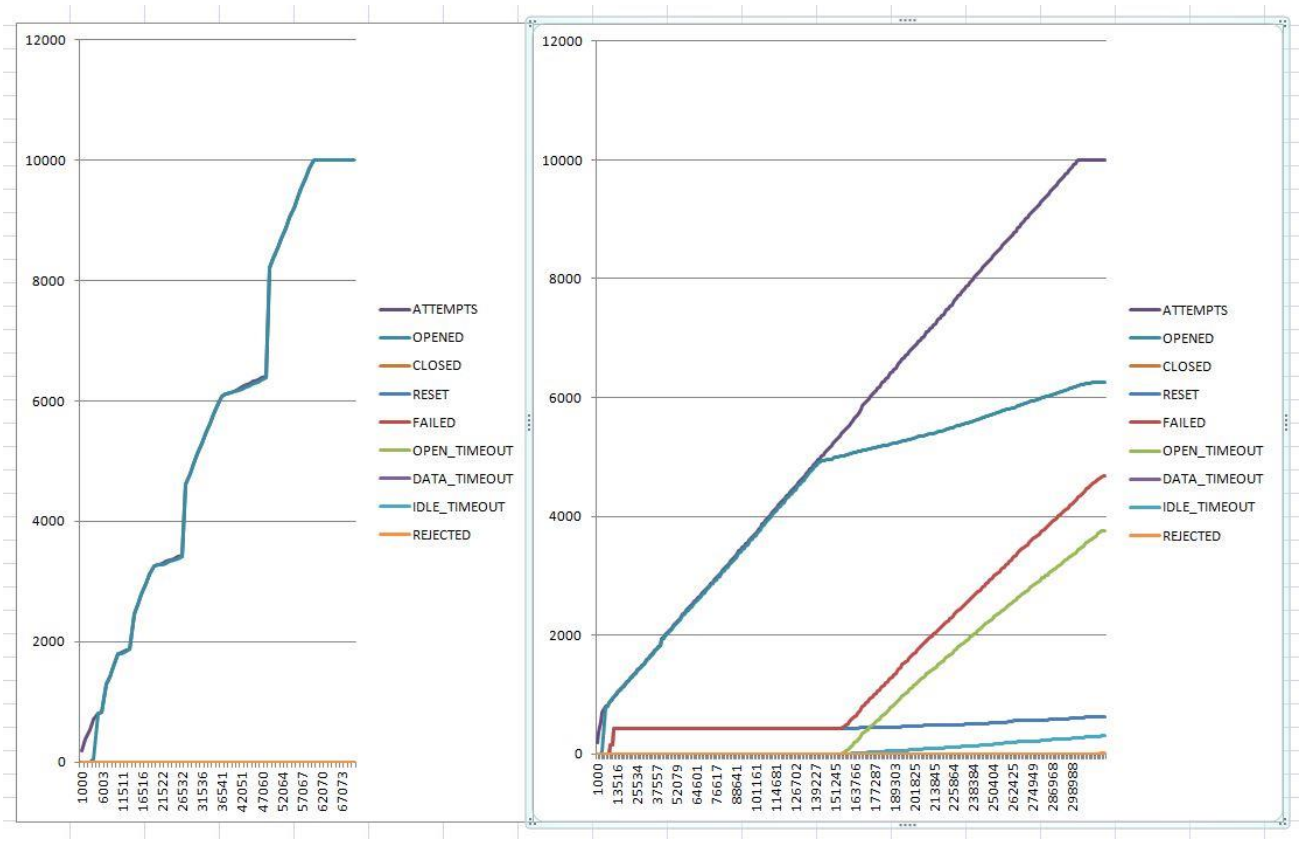


- Measured with Spec SFS2014
- Load = number of virtual clients from two Windows2012 machines

- Samba - degradation starts after 80 loads. It happens due to overload of either CPU or disk.
- NQ Storage – degradation starts somewhere after 300 loads (we did not test that far).
- No denial of service.

The later degradation happens, the more scalable SMB server is.

The results (high load)



- Measured with Load Dynamics
- Load = number of concurrent client connections

The solution with more resource consumption show denial of service.

Platform-wise scalability

Hardware range

Typical server platforms:

- SOHO NAS: ARM 1.2GHz Dual Core, HDD
- Mid-level storage: Atom® 2.13GHz Quad Core, HDD
- Top-end storage: Intel® Xeon® 3.4GHz Quad Core/IBM Power8, SSD.

Operating systems

- Some Storage vendors use VxWorks – an excellent platform for Storage because of a low-cost context switch. It is not freeware, though.
- Most of Storage vendors use Linux or (a kind of) Unix.
- Many of them, especially when it comes to Clustering, create their own software environment. Zero-latency networking (like DPDK), distributed FS, etc., add layers above layers on top of Linux/Unix. From the integration point of view such a platform can be considered as yet another OS.

A platform-wise scalable SMB must adapt to and run effectively on each of those platforms.

Measuring scalability

Tools matter

- Is SMB implementation A better than SMB implementation B? *This question is incorrect since the measuring tool matters.*
- A can be better than B when measured by SQLIO. With IOMETER the situation can be opposite.

Environment matters

- HDD vs SSD.
- Slow single-core CPU vs fast multi-core CPU with hyper-threading.

Methods matter

- A is better than B on 4K-packet random reads.
- B is better than A on a 64K-packet mix of 25% random writes.

Test tools

Tool	Method	Loads	Notes	Scalability
SpecSFS 2014	Load generator	•Desktop •Streaming •Compilation •Database	Only generic results, Very high loads. (in terms of traffic)	Usage-wise Load-wise
I/O meter	Read/write load	•Database •Streaming •Mix	Drill-down results	Usage-wise
SQLIO	SQL Server simulation	•Database	Basic benchmarking only	Basic
LoadDynamics	Hardware simulator	•Any + multiple connections	Very high loads (also in terms of connections)	Load-wise

Test tools (cont.)

- Platform-wise scalability – any tool goes
- “Nobody's perfect!..” – there is no single tool that measures everything
- Our experience - each customer has his favorite tool and methods

Conclusions

Processes vs threads

- The comparisons in this presentation were mostly done between NQ Storage 1.0 and Samba 4.3.4.
- Samba is process-based. Each client connection forks a new process. *100 connection == 100 processes.*
- NQ Storage is thread based. We use thread pool in the transport module and a number of thread pools in the SMB semantics module. On a low/moderate load this means that NQ attempts to allocate a thread for each request. *Threads are pre-allocated and their number (as well as resource consumption) is permanent.*

Processes vs threads (cont.)

- A process-based solution can introduce slightly better performance on a single SMB connection just because memory management (malloc) does not require locking. *This advantage seems to be negligible when pre-allocation is forced to avoid fragmentation.*
- A thread-based solution is more scalable. *This is in terms of load-wise, usage-wise and platform-wise scalability.*

Your feedback is very important for us.

Thank you

markr@visualitynq.com