ESG Lab Review

Performance Analysis: Scale-Out File Server Cluster with Windows Server 2012 R2

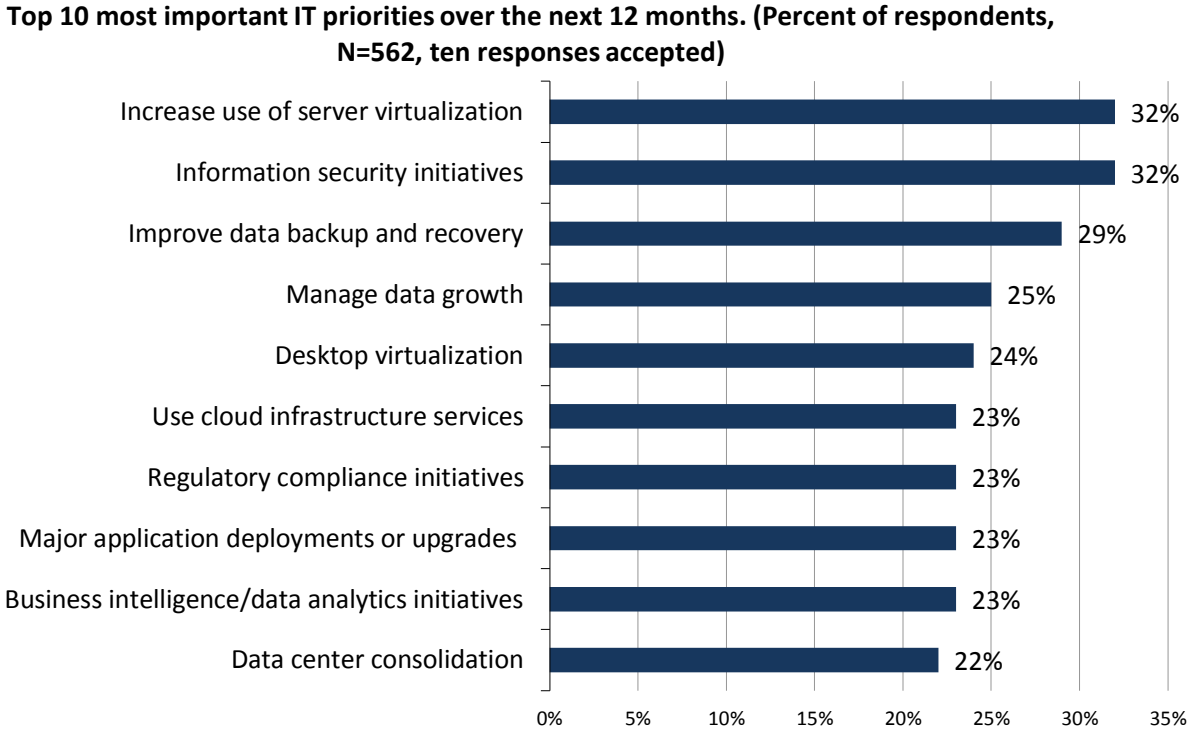
Date: December 2014 **Author:** Mike Leone, ESG Lab Analyst

Abstract: *This ESG Lab review documents the storage and virtualization performance of Hyper-V virtual machines running on a Windows Server 2012 R2 Scale-Out File Server (SOFS) cluster.*

The Challenges

Virtualization technology has helped reduce the cost and complexity of IT infrastructures for both large and small organizations around the globe, but challenges still exist. IT administrators continue to face pressure to improve infrastructure efficiency, performance, and availability, while future-proofing hardware investments to address the never-ending growth of data. In fact, in ESG's *2014 IT Spending Intentions Survey*, both the increased use of server virtualization (32%) and managing data growth (25%) appeared in the top five most-cited important IT priorities over the next 12 months.[1]

Figure 1. Top Ten Most Important IT Priorities



Source: Enterprise Strategy Group, 2014.

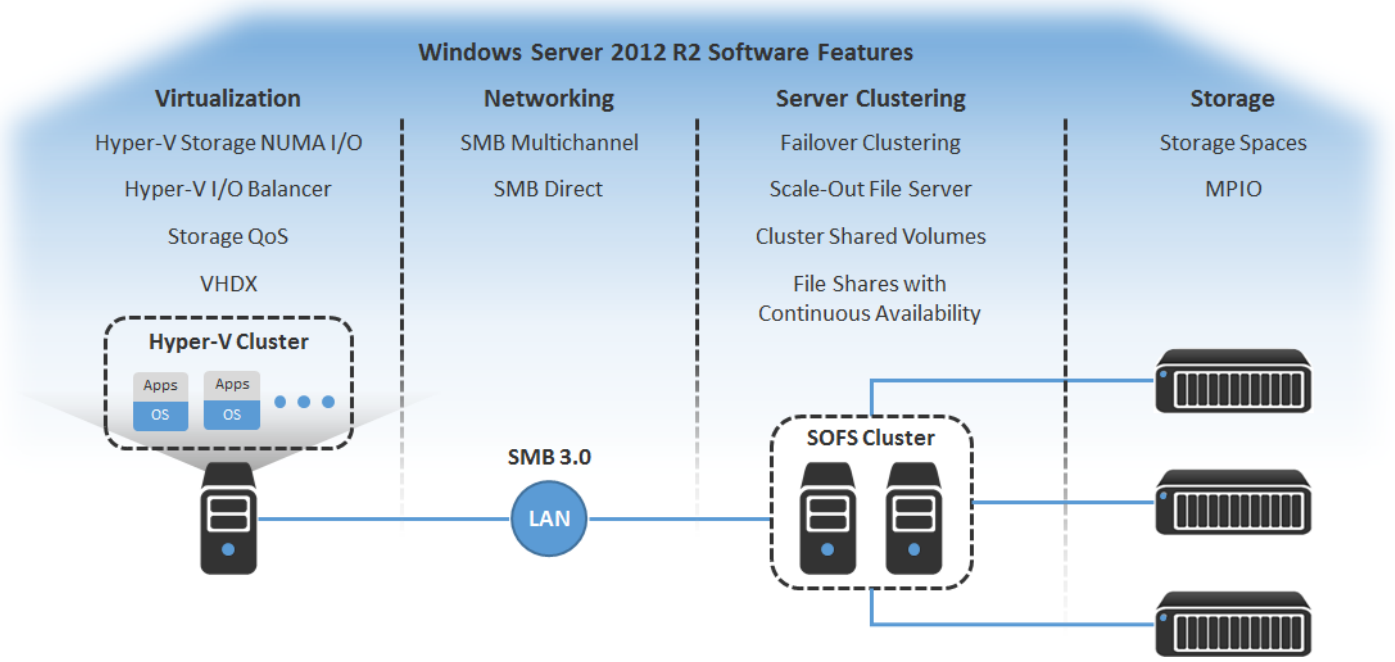
Next comes the budget restrictions: IT administrators are constantly faced with scenarios where hardware or software tradeoffs must be made because of cost. This situation puts a big focus on solutions that can reduce both CapEx (acquisition) and OpEx (management) costs while improving return on investment (ROI). This lends itself to the fact that the top measure organizations are taking to reduce and contain IT expenditures in 2014 is purchasing new technologies with improved ROI.[2] Until recently, delivering enterprise-class performance, resiliency, availability, and flexibility on a tight budget was a pipe dream, but with Windows Server 2012 R2 it is now becoming a reality.

[1] Source: ESG Research Report, *2014 IT Spending Intentions Survey*, February 2014.
[2] Ibid.

Windows Server 2012 R2: Storage and Virtualization Software Features

Windows Server 2012 R2 provides powerful technologies that help transform an existing IT infrastructure into a modern data center that effectively addresses every available opportunity with agility and efficiency. New and improved features provide high performance and scalability that create a dynamic, multitenant infrastructure, which securely scales workloads to meet even the most complex SLAs, while offering the flexibility to adapt to ever-changing business needs. Microsoft has a comprehensive set of enterprise-class storage and virtualization features delivered on cost-effective hardware to meet the demands of even the most cost-conscious customer. These important features are shown in Figure 2 and can be separated into four key areas: virtualization with Hyper-V, networking with SMB 3.0, server clustering with Scale-out File Server (SOFS), and storage.

Figure 2. Windows Server 2012 R2 Storage and Virtualization Software Features



Virtualization with Microsoft Hyper-V

Hyper-V in Windows Server 2012 R2 provides organizations with massive scalability to transform data centers into a complete cloud platform. Hyper-V greatly expands support for resource allocation, helping to ensure that any virtualization infrastructure can support the configuration of large, high-performing virtual machines to maintain workloads that require significant scale. Microsoft designed Windows Server 2012 R2 with Hyper-V to meet data center scalability requirements in order to allow for improved virtual machine density and higher performing servers with hardware innovations and acceleration technology.

With *Storage NUMA I/O*, communication channels are created between the host storage stack and guest devices. Each guest device gets assigned a dedicated set of virtual processors that are used to process storage I/O for that particular device. This creates improved efficiencies for I/O completion and eliminates inter-processor interruptions. This is particularly important for large virtual machines that can allocate up to 64 virtual processors with the latest version of Hyper-V. NUMA I/O ensures that these large, powerful VMs running mission critical applications create an optimal mapping between virtual and physical resources, making certain that guest operating systems and applications can achieve the most efficient execution, best performance, and most linear scale.

Virtual *Storage Quality of Service (QoS)* and *Hyper-V I/O Balancer* provide the capability to set I/O performance thresholds for individual virtual hard disks. This helps to prevent a single virtual machine from consuming too much storage performance, ensuring a predictable balance across all VMs in a Hyper-V cluster. And finally, the new *VHDX* format supports up to 64TB of virtual hard disk capacity with built-in resiliency and an additional 4KB logical sector to improve performance for certain workloads and applications.

Storage Networking with SMB 3.0

The SMB protocol is a network file sharing protocol that allows applications and IT end-users to access files or other resources from a remote file server, allowing applications to read, create, and update files on the remote file server. The SMB protocol can also communicate with any server program that is set up to receive an SMB client request. *SMB Multichannel* allows servers to use multiple network connections simultaneously to increase both performance and availability. Data is transmitted across multiple network connections on high-speed network adapters or across multiple network adapters to aggregate the performance. *SMB Direct* enables the use of remote direct memory access (RDMA) for supporting network adapters, which consumes fewer CPU cycles and lowers latency while increasing performance. This allows applications access to SMB-based storage shares at speeds that rival direct-attached storage. Currently, SMB Direct supports three types of RDMA: InfiniBand, Internet Wide Area RDMA Protocol (iWARP) and RDMA over Converged Ethernet (RoCE).

Server Clustering with Windows Server Scale-Out File Server cluster

A group of computers that work together for increased availability and scalability is a *failover cluster*. If one or more clustered nodes fail, other nodes will maintain availability without disrupting service. The *Scale-Out File Server (SOFS)* role is then layered on top of the failover cluster to provide always-available SMB service that takes advantage of the aggregate bandwidth across all the nodes in the cluster. The SOFS role enables SMB clients to get transparently directed to their "owner" node for resource balancing, which helps to improve performance. *Cluster Shared Volumes (CSV)* are used to give multiple nodes in the cluster simultaneous access to shared storage within a single, distributed namespace, simplifying management of the underlying storage volumes in a failover cluster. Finally, with file shares configured with *Continuous Availability (CA)* enabled, file handles are persistently left open with write-through to guarantee data integrity, stability, and durability against potential node failures within the cluster.

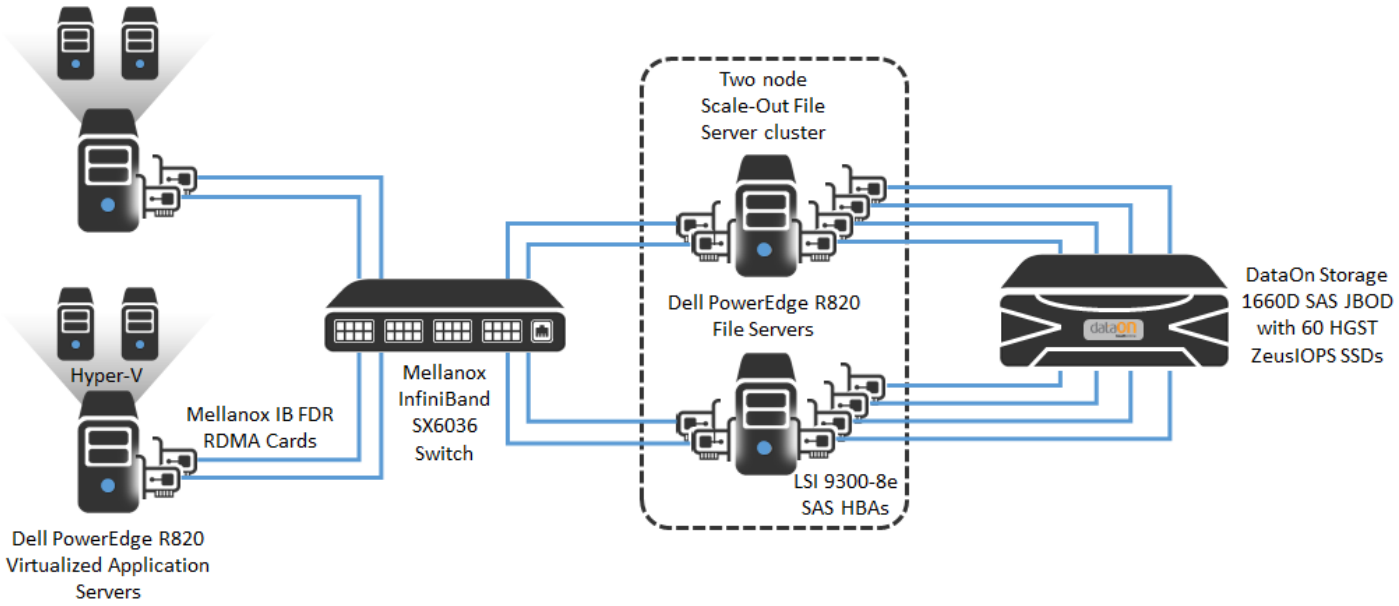
Virtualized Storage with Storage Spaces

Storage Spaces in Windows Server 2012 R2 is designed to allow for the creation of highly available, scalable, and performing storage solutions by virtualizing cost-effective storage. Storage Spaces is designed with key technologies to meet the needs of enterprise workloads. Storage pools aggregate sets of physical disks into one or more units of management that can be expanded dynamically. Storage Spaces provides storage virtualization and resiliency with support for clustering. Storage Spaces in Windows Server 2012 R2 provides the perfect foundation for SMB 3.0 file shares while offering a single point of management for storage, server, and network. With *Multipath I/O*, storage providers can leverage a framework built by Microsoft to develop multipath solutions that contain the hardware-specific information required to optimize connectivity with their particular storage arrays.

Getting Started

Testing began with a tour of the Microsoft-configured test bed (see Figure 3). Two Dell PowerEdge R820 servers running Windows Server 2012 R2 with Hyper-V were used as the SMB clients. Each SMB client had six VMs, each configured with 16 virtual processors with one virtual NUMA node, 16GB RAM, and one virtual SCSI controller with 16 VHDX files attached. Each VHDX file was 127GB fixed type and hosted on a separate SMB share in the Scale-Out File Server cluster. The two client servers were connected via two 56Gb Mellanox IB FDR RDMA cards per server to a 56Gb Mellanox InfiniBand SX6036 switch. Both ports of each adapter were used during testing. The Mellanox switch was then connected to a two-node SOFS Cluster, made up of two additional Dell PowerEdge R820 file servers. Two of the six PCIe slots in each file server node contained two 56Gb Mellanox IB FDR RDMA cards, while the other four PCIe slots contained 12Gb LSI 9300-8e SAS HBAs. Both sets of four HBAs were connected to both storage controller heads of a DataOn 1660D SAS JBOD Storage array. This dual-level connectivity ensured high levels of both performance and redundancy. The 60-bay storage system was filled with 60 300GB HGST ZeusIOPS SSDs. The storage was divided into eight virtual disks (for each VM) and backed by Storage Spaces.

Figure 3. Performance Test Bed

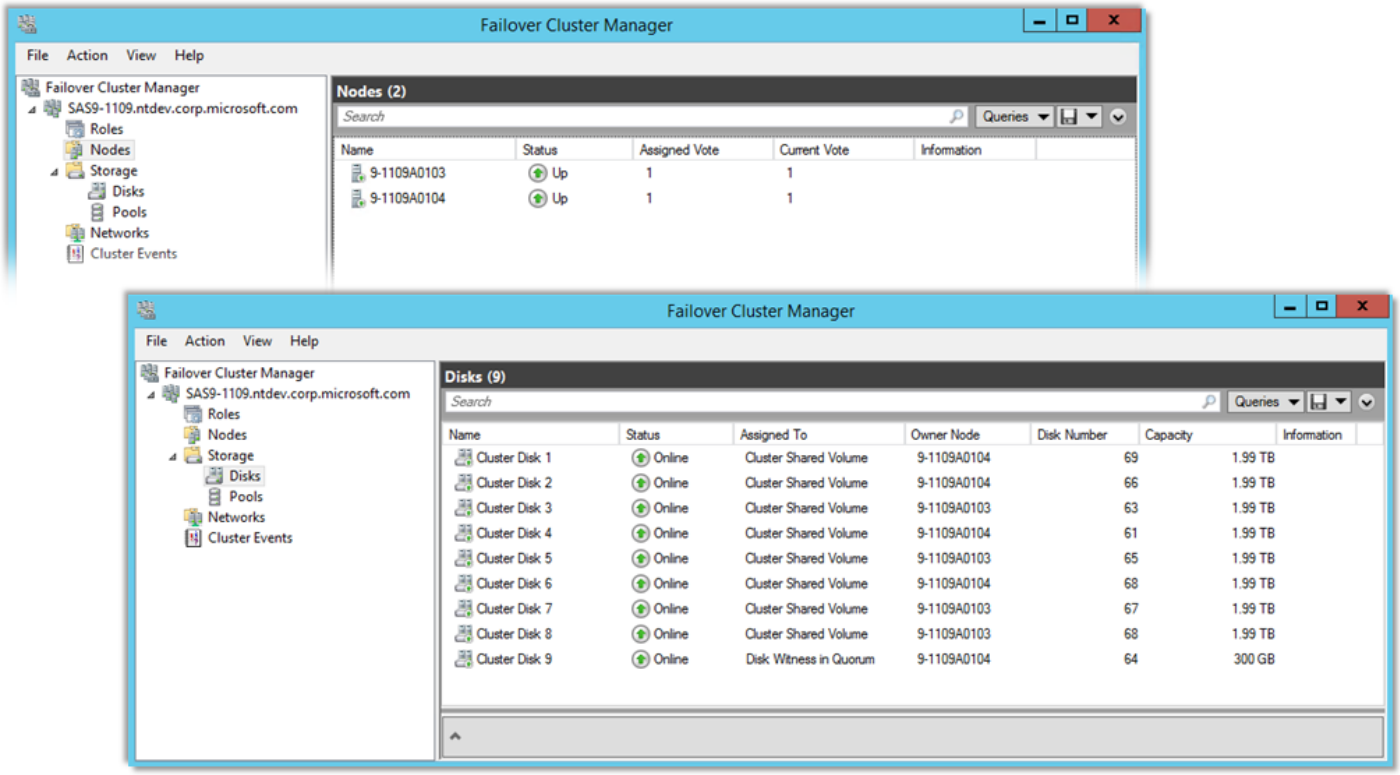


Multiple steps were required from a server, storage, and network standpoint to enable a highly performing Scale-Out File Server platform. From a server standpoint, the SOFS is built on top of a Windows Server Failover Cluster. The top picture in Figure 4 shows a snapshot of Failover Cluster Manager where a two-node failover cluster was created. Next, the SOFS role was added to the failover cluster, which finished the cluster creation.

Once the two-node SOFS was created, it was time to configure the storage. A single storage pool striped across all 60 SSDs was created and then eight simple Storage Spaces were created and added to the SOFS cluster (see bottom picture in Figure 4). Cluster Shared Volumes were added to each shared disk, which serve as distributed file system access systems that enable every node in the cluster to concurrently access a shared volume. This unifies the storage access into a single namespace for easier management and improved VM migration capabilities. A total of 128 SMB file shares with Continuous Availability (CA) were created in the SOFS cluster, which enabled seamless service on any node in the cluster without interrupting the server applications, while also facilitating better load balancing.

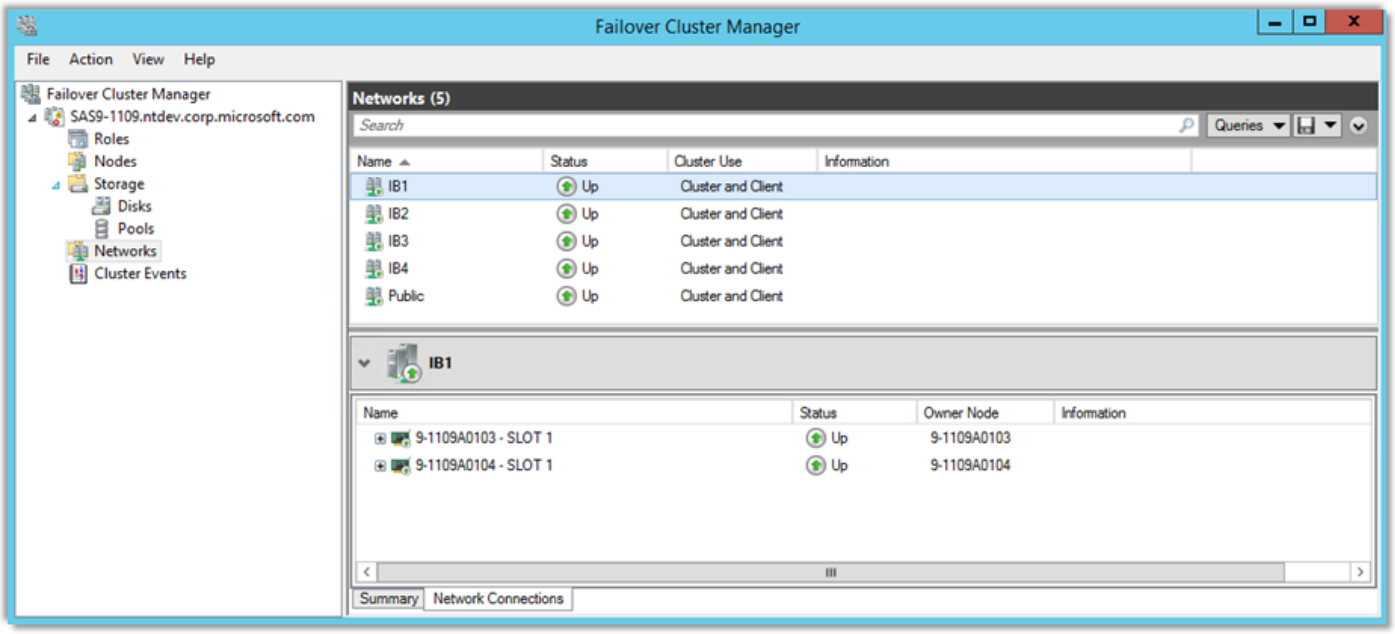
A ninth simple Storage Space was also added to the SOFS cluster to serve as the cluster quorum witness service. The quorum for a cluster is the number of elements that must be online for that cluster to continue running. Each element can cast one "vote" to determine whether the cluster continues running. The voting elements are nodes, a disk witness, or a file share witness. Each voting element (with the exception of a file share witness) contains a copy of the cluster configuration, and the cluster service works to keep all copies synchronized at all times. If too many failures occur in the cluster, the quorum service will bring down the entire cluster to avoid data loss or data inconsistency.

Figure 4. Enabling SOFS Cluster Role and Creating Cluster Shared Volumes



Finally, from a networking standpoint, SMB Multichannel was configured to allow for the use of multiple paths simultaneously. For the SMB Multichannel to work properly, a separate subnet must be configured for every NIC in the cluster because SOFS clustering will only use one IP address per subnet regardless of the number of NICs on that subnet. Four subnets were used in this cluster, which were dedicated to each of the four InfiniBand connections. As shown in Figure 5, each InfiniBand network connection was assigned to allow both cluster network communication and client traffic, while the InfiniBand network connections are owned by both nodes in the SOFS cluster.

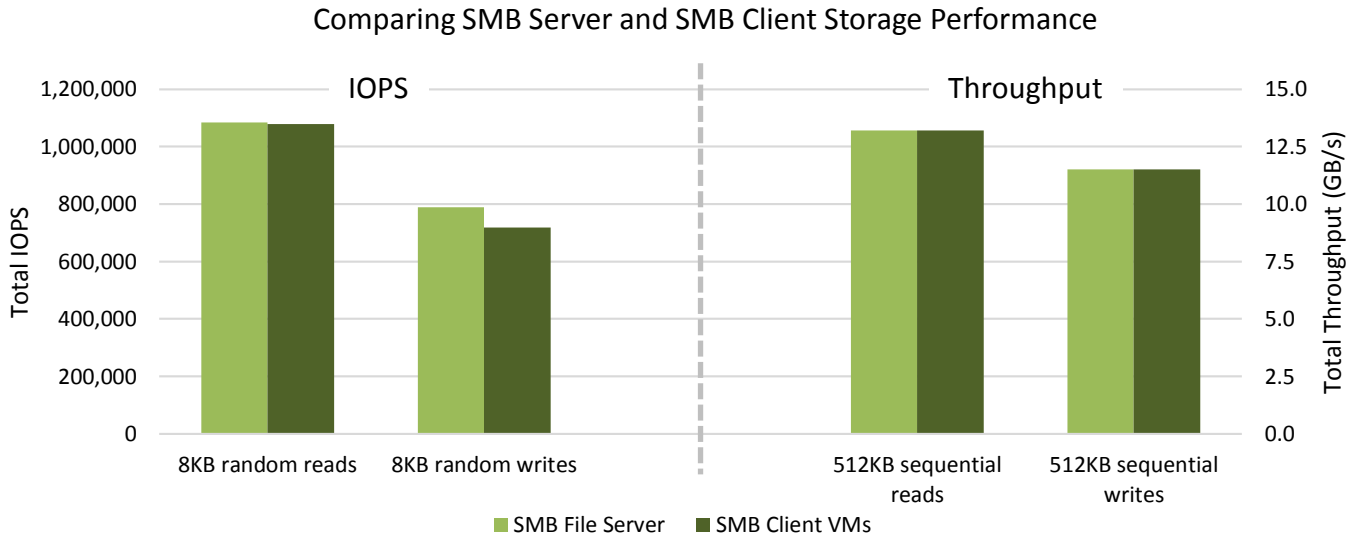
Figure 5. Configuring SMB Multichannel



Performance Analysis

This section of the report presents the ESG Lab-audited performance results with a goal of showing the performance capabilities of leveraging Windows Server 2012 R2 storage and virtualizations technologies. The industry-standard Iometer tool was used as the storage workload generator to test read, write, and mixed workload performance. The first phase of testing focused on the pure read and write workloads with a goal of achieving high levels of storage performance. Two test scenarios were used during this phase of testing to highlight the efficiency of the SMB 3.0 protocol. The first scenario consisted of running Iometer on the file servers with local Cluster Shared Volumes, while the second scenario measured performance from within the hosted virtual machines on the two SMB clients. IOPS was measured for random 8KB reads and writes, and throughput was measured for sequential 512KB reads and writes. For IOPS measurements, four total client VMs were used, while throughput measurements utilized a total of eight client VMs. The results are shown in Figure 6.

Figure 6. SMB 3.0 Efficiency for IOPS and Throughput



What the Numbers Mean

- The 8KB random read workload yielded similarly high levels of performance, achieving over 1,000,000 IOPS for both server- and client-side scenarios. The SMB file server test achieved a maximum of 1,084,000 IOPS, while the SMB client VMs yielded a less than 1% difference at 1,078,000 IOPS.
- CPU utilization was monitored during the peak IOPS test runs. The SMB File Server tests yielded 60% CPU utilization, while the SMB client VMs utilized 40% of the CPU, reducing CPU utilization by over 30%.
- For 8KB random writes, performance was equally impressive, with server-side IOPS reaching a peak of 789,000 IOPS, while client side achieved a maximum of 719,000—a manageable difference of under 10%.
- For throughput testing, 512KB sequential reads and writes showed no difference when comparing server side with client side. Read throughput reached over 13 GB/s, while write throughput peaked at 11.5 GB/s.
- In all test cases, ESG Lab came to the same conclusion: The SMB 3.0 protocol had little impact on overall storage performance. This proved from a pure IOPS or throughput standpoint that the SMB 3.0 protocol was more than efficient enough to handle these high levels of performance using a SOFS platform built with industry-standard storage hardware.

The next phase of testing simulated two important business application workloads: OLTP Database and Exchange Server. Like most business applications, these workloads had both read and write I/Os. The OLTP Database simulation consisted of a 100% random, 8KB workload with 90% reads and 10% writes, while the Exchange Server consisted of an 80% random, 20% sequential, 32KB workload with 60% reads and 40% writes. The workloads were run from both SMB clients within the configured virtual machines and IOPS, throughput, and response times were collected with a goal of showing the even workload distribution between SMB clients. The IOPS results from the OLTP testing and average response times from the Exchange Server testing are shown in Figure 7. All mixed workload application simulation test results are shown in Table 1.

Figure 7. SMB Client Performance Results for Simulated Business Applications

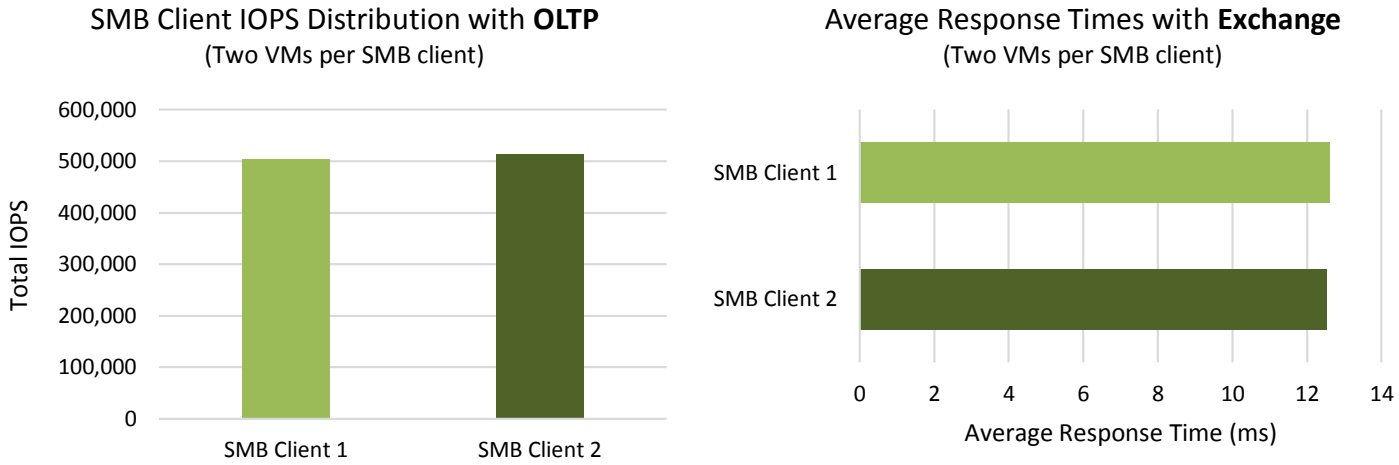


Table 1. SMB Client Performance Results for Simulated Business Applications

Client	OLTP Database			Exchange Server		
	IOPS	Throughput (GB/s)	Response Time (ms)	IOPS	Throughput (GB/s)	Response Time (ms)
Client 1	504,000	3.9	4.06	161,000	5.05	12.62
Client 2	513,000	4.0	3.99	162,000	5.09	12.54
Total	1,017,000	7.9		323,000	10.14	

What the Numbers Mean

- In both enterprise-class application simulations, the workloads yielded high levels of predictable performance that were evenly distributed across both SMB client servers.
- The OLTP database workload achieved over 1,000,000 IOPS with over 500,000 IOPS coming from each SMB client. Average response times remained predictably low, with a maximum of 4.06ms and a minimum of 3.99ms, while aggregate throughput achieved nearly 8 GB/s.
- For the more complex Exchange Server workload, both random and sequential read and write I/Os produced approximately 323,000 aggregate IOPS. Response time, being an important measurement of a successful Exchange deployment, peaked at 12.62ms, well under Microsoft's recommended thresholds, while aggregate throughput performance achieved over 10 GB/s.

Why This Matters

Create a cost-effective, easy-to-manage solution that delivers high levels of performance for virtualized environments, while offering enterprise-class reliability and scalability—sounds easy enough, right? While meeting all of these requirements sounds achievable with a number of solutions in the market today, the price point is where the conversation tends to go sideways. Because of cost, tradeoffs must be made that sacrifice important must-haves like performance or manageability, eventually leading to future growth problems.

ESG Lab validated that a cost-effective Microsoft Scale-out File Server Platform that uses the latest Windows Server 2012 R2 storage and virtualization technologies can achieve high levels of enterprise-class storage performance. SMB 3.0 efficiency was proven by comparing SMB server- and client-side storage performance, yielding minimal to no impact on IOPS and throughput. Simulated enterprise applications like OLTP and Exchange Server yielded predictable performance across SMB client nodes. ESG Lab was particularly impressed with the client-side peak performance results—**up to 1.07 million IOPS and 13.2 GB/s of throughput**.

The Bigger Truth

Server virtualization has had a major impact on IT organizations of all sizes, helping to deliver significant business value, lower costs, and improve resource utilization. While these desired benefits are significant, they don't come without their challenges, especially when it comes to the impact virtualization has on a storage infrastructure. As organizations grow, storage requirements become not only harder to manage, but also harder to achieve, especially in virtualized infrastructures running enterprise applications with strict performance requirements. Being able to address these challenges cost effectively is an additional challenge in itself. Vendor-specific hardware, vendor-specific expertise for management purposes, costly upgrades, never-ending service contracts, continued licensing, and ongoing training all add up quickly.

When Microsoft first introduced Windows Server 2012, new and improved technologies were launched that addressed many storage and virtualization challenges. This trend continues with the release of Windows Server 2012 R2, with even more features and functions being added that enhance the already robust feature set. Now organizations have a solution that addresses not only the performance, high-availability, reliability, and management challenges faced by many continuously growing virtualized IT environments, but also the cost challenges. Hyper-V virtualized applications can leverage the Scale-Out File Server platform on Windows Server 2012 R2 to achieve high levels of storage performance and resiliency similar to a traditional SAN environment. This is done with a Failover Cluster, running the Scale-Out File Server role, backed by resilient Storage Spaces, delivering virtualized storage, connected via high-performance SMB 3.0 networking, all running on cost-effective, industry-standard hardware. ESG Lab confirmed that Hyper-V VMs in a two-node Scale-Out File Server cluster running on Storage Spaces, built on industry-standard hardware, achieved very high levels of storage performance. Up to 1.1 million IOPS and 13 GB/s were measured for pure read and write workloads. OLTP Database and Exchange Server simulations also achieved impressive levels of performance. Virtualized OLTP workloads yielded over 1 million IOPS and virtualized Exchange Server workloads exceeded IOPS and throughput expectations, while meeting low response time requirements.

Simplicity, manageability, efficiency, flexibility, scalability, resiliency, high availability, performance, cost-effectiveness — Microsoft storage and virtualization technology packs a major punch. As the robust feature set continues to expand and mature, it's no surprise that organizations are adopting the technology now to improve and, in some cases, replace their existing infrastructure platforms. If strict budget requirements are forcing your organization to make difficult IT investment decisions, ESG Lab suggests taking a look at Microsoft's latest storage and virtualization technology.

For further technical depth on the configuration and performance tests, please refer to the accompanying Microsoft whitepaper, entitled "Achieving over 1-Million IOPS from Hyper-V VMs in a Scale-Out File Server Cluster using Windows Server 2012 R2" co-authored by Liang Yang, Danyu Zhu, Jeff Woolsey and Senthil Rajaram. This whitepaper can be downloaded from http://www.microsoft.com/en-us/download/details.aspx?id=42960.