



Microsoft[®] Exchange Server Performance on VMware Virtual SAN[™]

Performance Study

TECHNICAL WHITE PAPER

Table of Contents

Executive Summary	3
Introduction.....	3
Experimental Configuration and Methodology	3
Test-Bed Configuration	3
Test and Measurement Tools.....	4
Test Cases and Test Method	5
Experimental Results and Performance Analysis.....	6
Conclusion	6
Appendix A. System Under Test.....	7
Compute/Storage hosts.....	7
LoadGen host	7
References	8

Executive Summary

VMware® performance testing shows that Microsoft® Exchange Server 2010 on Virtual SAN clusters scales well without affecting much user-perceived application latency as more Exchange users are deployed with added VMware Virtual SAN™ hosts.

Introduction

Microsoft Exchange Server is a commonly found email server and is considered a business-critical application by many organizations. Virtualized instances of Exchange Server can be successfully deployed using VMware vSphere® 5.5 [1] and it has been shown that Exchange Server performs well in a virtualized environment with shared SAN storage [2]. With the release of VMware Virtual SAN [3], the next logical step is to study the performance of Exchange Server on this storage platform.

This paper shows that virtualized Exchange Server 2010 on vSphere 5.5 with VMware Virtual SAN scales well while maintaining very good application response time as more Exchange users are added along with more Virtual SAN hosts. This scale-out approach of supporting more Exchange users by adding more Virtual SAN hosts as needed is a very cost-effective way to accommodate growing email users without compromising performance.

VMware Virtual SAN is a software-defined storage layer that runs natively on the hypervisor in vSphere hosts. By clustering server solid-state drives (SSD) and hard disk drives (HDD) across vSphere hosts, Virtual SAN provides highly resilient and shared storage for vSphere clusters. The Virtual SAN distributed architecture leverages enterprise grade SSDs for high performance read/write caching and HDDs for cost-effective data persistence. This paper demonstrates that the Virtual SAN SSD caching layer handles incoming read/write I/Os from Exchange Mailbox virtual machines without affecting application latency as more Exchange users are added.

Experimental Configuration and Methodology

The performance and sizing studies were done in VMware's internal labs. The purpose of the tests was to measure, analyze, and understand the performance of Exchange on a VMware Virtual SAN environment.

Test-Bed Configuration

Servers used for Exchange testing were configured based on sizing best practices. Five Dell PowerEdge R720xd servers were configured with two Intel® Xeon® Processors E5-2650 and 128GB of physical memory as hosts for Virtual SAN clusters as shown in [Figure 1](#). Exchange Server 2010 was installed on Windows Server 2008 R2 in a virtual machine. The virtual environment used VMware vSphere 5.5 U1 to virtualize the Windows/Exchange environment. There are several Exchange server roles that are required for a functional Exchange environment as well as the requirement of a Windows Active Directory with DNS. Virtual SAN was used to provide the shared storage for the Exchange Mailbox, CAS-HUB virtual machines and the Windows Active Directory/DNS virtual machine. In addition to the Virtual SAN hosts, one Dell PowerEdge R720xd server was configured with two Intel Xeon Processors E5-2690 and 128GB of physical memory for running the client virtual machine with the Exchange Load Generator test workload. The BIOS power profile of the servers was set to maximum performance.

For additional test-bed configuration, see [Appendix A](#).

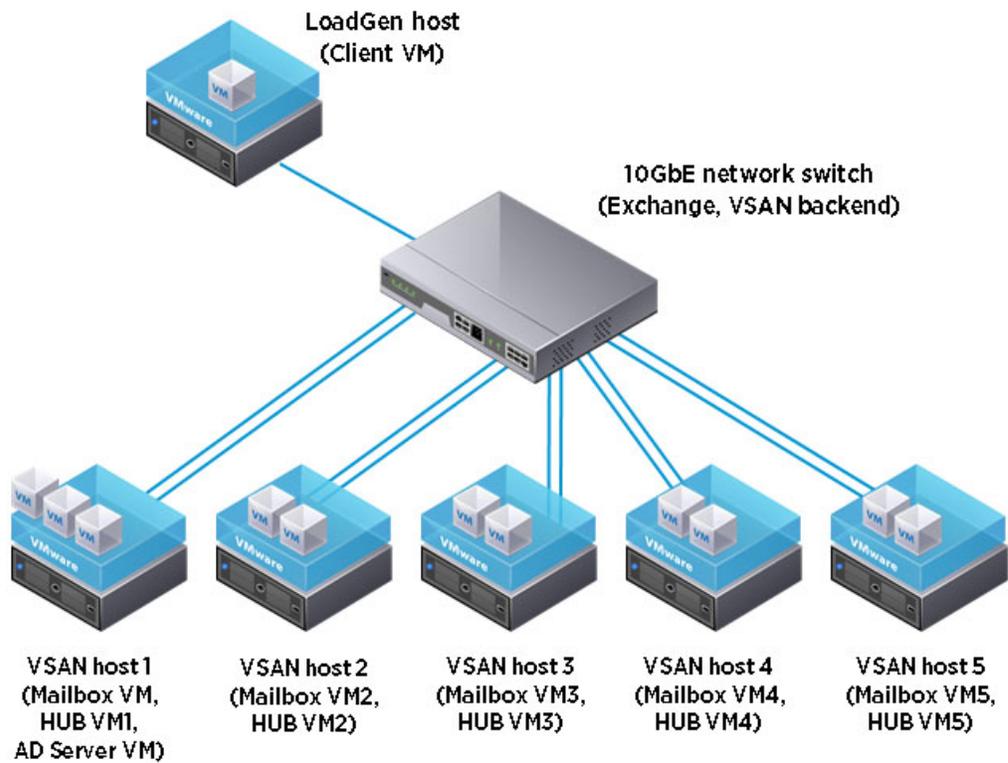


Figure 1. Test-bed layout

Test and Measurement Tools

Microsoft's Exchange Load Generator 2010 (LoadGen) runs on a client system and simulates the messaging load against Exchange deployments. LoadGen measures user response times for each of the different types of operations that it runs against the Exchange systems. In this case, LoadGen was configured to simulate Outlook 2007 online clients using the LoadGen very heavy user profile with 150 messages sent and received per day, per user. Each mailbox user was initialized with 100MB of data in the mailbox before starting the tests. LoadGen tests include a variety of Exchange transactions and were run for an 8 hour period to represent a typical workday. Quality of Service (QoS) is determined by Industry consensus, which indicates that testing should focus on the 95th-percentile transaction latency of the Sendmail transaction to measure mailbox performance. The value of this result, which represents the maximum time needed to send an email for 95% of the transactions, should be reported below 500 milliseconds (ms) to represent an acceptable user experience.

The LoadGen benchmark tries to closely model the normal daily email usage of real users, in order to provide an estimate of the number of users a system can support. While this benchmark is widely used to measure the performance of Exchange platforms, as with all benchmarks, the results may not match the specifics of your environment.

Microsoft Windows Perfmon [4] was used to collect the latency and status counters of Exchange client tasks in the client machine.

For vSphere 5.5, esxtop was used to record both vSphere and virtual machine-related performance counters. Esxtop was configured to log CPU, memory, disk, network, and system counters during LoadGen runs.

Test Cases and Test Method

The configuration of user and resource (processor and memory) sizing was based on Microsoft's recommendations [5] [6]. 1,000 very heavy users were configured per processor and memory was sized to 4GB plus 6MB per mailbox for the Exchange mailbox virtual machine.

In this paper, the scale-out performance of Exchange on Virtual SAN is shown by increasing the number of virtual machines. The Exchange environment was scaled out by deploying one combined Client Access/Hub Transport virtual machine for every mailbox server virtual machine in a 1:1 ratio as recommended in the "Performance and Scalability" section for Exchange 2010 on Microsoft TechNet [7].

The two types of virtual machine configurations for the scale-out tests were:

- Mailbox server virtual machine with 4 virtual CPUs and 28GB memory to support 4,000 users.
- Client Access and Hub Transport combined server virtual machines with 4 vCPUs and 8GB memory.

To demonstrate scale-out performance, the number of Exchange virtual machines was increased to ten (five mailbox server virtual machines + five Client Access and Hub Transport combined role VMs) with 20,000 Exchange very heavy users. The first Virtual SAN host ran an AD/DNS virtual machine in addition to the mailbox and CAS-HUB virtual machines.

Table 1 shows the number of CPUs, memory size, and number of very heavy users for each scenario in the scale-out experiments.

Number of mailbox and CAS/HT VM	Total Number of vCPUs in VMs	Total Memory (GB) in VMs	Total Number of Very Heavy Users
3+3	24	108	12,000
4+4	32	144	16,000
5+5	40	180	20,000

Table 1. Number of VMs, vCPUs, memory, and users in scale-out scenarios

In each of the scalability test scenarios, the Exchange transaction latencies were used to measure performance. Transaction latency directly predicts user experience. The tests were run for four hours and five times of each test to get the average from three median results.

Experimental Results and Performance Analysis

Among all of LoadGen's reported transaction latencies, Sendmail is the most prominent and its latency is considered a yardstick of user responsiveness. The 95th-percentile for Sendmail latencies being at or below 500ms is considered acceptable user responsiveness. Sendmail is therefore used to compare performance across the various configurations.

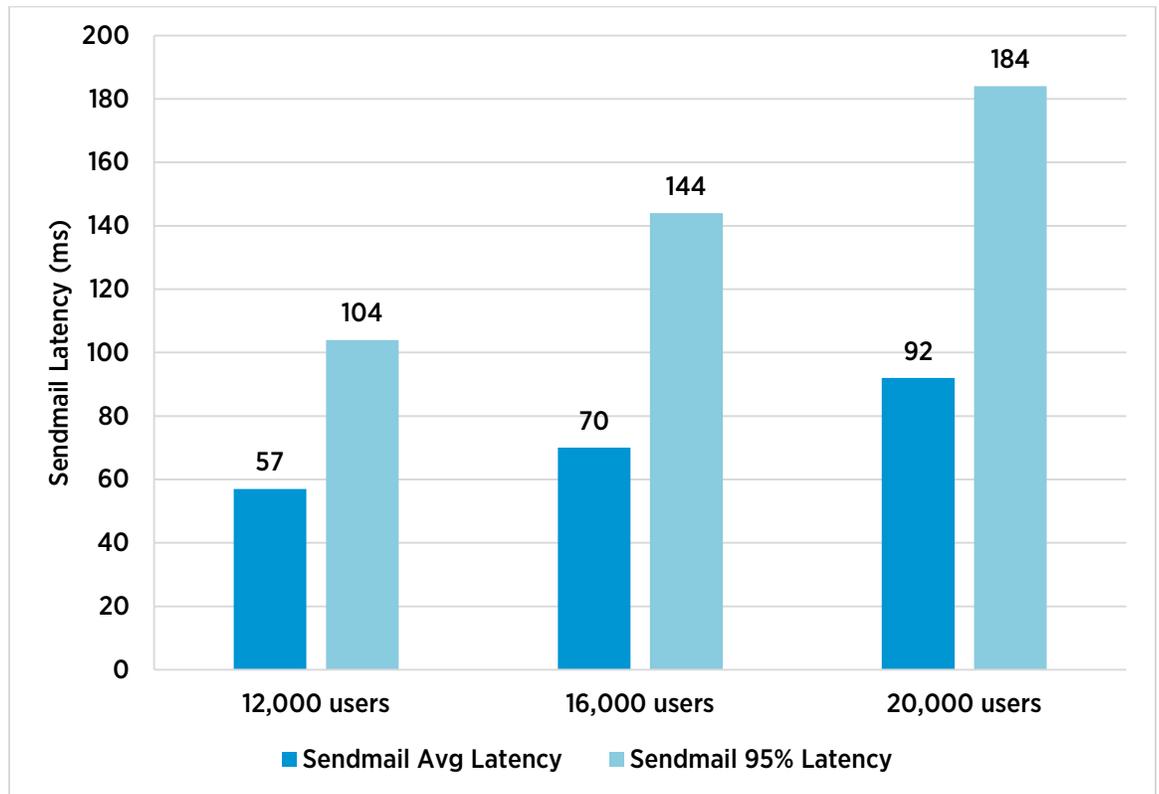


Figure 2. Average and 95th Percentile Sendmail Action Latency

Figure 2 shows the average Sendmail latency is only 57ms with a user load of 12,000 and is increased up to 92ms even with 20,000 users. The 95th-percentile Sendmail latency of the 12,000 user load case is 104ms, which is much lower than the standard maximum of 500ms. Even the 95th-percentile Sendmail latency of the 20,000 user load case is 184ms, which is still far lower than 500ms. This result shows that Virtual SAN serves as an excellent distributed and scalable-shared storage for an Exchange workload.

Conclusion

The results in this paper show that Exchange 2010 running on Virtual SAN achieves very good performance and scalability. The paper shows that the Exchange Sendmail 95th-percentile latencies are far better than the standard requirement of 500ms from a load of 12,000 users to 20,000 users.

These low latencies reported for 95% of the mailboxes provide an outstanding response time for Exchange users. The performance results and analysis demonstrates that VMware Virtual SAN is an ideal scale-out storage solution for growing Exchange deployments.

Appendix A. System Under Test

The following lists show how the Virtual SAN hosts and load generator client were configured as the system under test (SUT).

Compute/Storage hosts

- **Servers:** Dell PowerEdge R720xd
- **Processors:** 2 Intel Xeon E5-2650 processors @ 2.00GHz, Hyperthreading enabled
- **Memory:** 128GB Memory
- **Storage HBA:** LSI SAS 9208 8i Adapter
- **Network Adapters:** Intel X520-SR2 two port 10Gb Ethernet Server Adapter
- **Virtual SAN datastore:** 3 disk groups per host; each disk group consists of 1 SSD (Toshiba 200GB SLC 6Gbps SAS) and 4 HDDs (Seagate 900GB 6Gbps SAS)
- **Virtual machines** (Mailbox, Active Directory, DNS, Exchange Client Access and Hub Transport roles):
 - **Mailbox VM:** 4 vCPUs, 28GB memory, 35GB OS vmdk, 2x400GB mailbox vmdks, 1x800GB backup vmdks, 200GB log vmdk
 - **CAS-HUB VM:** 4 vCPUs, 8GB memory, 35GB OS vmdk
 - **AD/DNS VM:** 8 vCPUs, 32GB memory, 35GB OS vmdk
 - One Mailbox VM and one CAS-HUB VM per a host
 - One AD/DNS VM per clusterMailbox VM had 4000 users, very heavy profile

LoadGen host

- **Client:** Dell PowerEdge R720xd
- **Processors:** 2 Intel Xeon E5-2690 processors @ 2.90GHz, Hyper-Threading enabled
- **Memory:** 128GB Memory
- **Storage HBA:** LSI SAS 9208 8i Adapter
- **Network Adapters:** Intel X520-SR2 two port 10Gb Ethernet Server Adapter
- **VMFS datastore:** 1 HDD (Seagate 900GB 6Gbps SAS)
- **Virtual machine:**
 - Client VM: 8 vCPUs, 32GB memory, 35GB OS vmdk

References

- [1] VMware, Inc. (2014) Virtualizing Exchange with VMware.
<http://www.vmware.com/business-critical-apps/exchange/index.html>
- [2] VMware, Inc. (2011) Microsoft Exchange Server 2010 Performance on VMware vSphere 5.
<http://www.vmware.com/files/pdf/exchange-perf-vsphere5.pdf>
- [3] VMware, Inc. VMware Virtual SAN.
http://www.vmware.com/files/pdf/products/vsan/VMware_Virtual_SAN_Datasheet.pdf
- [4] Microsoft. Perfmon. <http://technet.microsoft.com/en-us/library/bb490957.aspx>
- [5] Microsoft. (2011, February) Understanding Processor Configurations and Exchange Performance.
<http://technet.microsoft.com/en-us/library/dd346699.aspx>
- [6] Microsoft. (2012, April) Understanding Memory Configurations and Exchange Performance.
<http://technet.microsoft.com/en-us/library/dd346700.aspx>
- [7] Microsoft. (2010, September) Understanding Client Access and Hub Transport Combined Role Configurations in Capacity Planning. <http://technet.microsoft.com/en-us/library/ee832795.aspx>

About the Authors

Jinpyo Kim is a staff engineer in the Performance Engineering team at VMware. He has worked on optimizing the performance of ESXi storage related products including SW iSCSI, PVSCSI, core storage stack, Storage I/O Control, Storage DRS, and VSAN. He received a PhD in Computer Science from the University of Minnesota-Twin Cities.

Tuyet Pham is a senior member of technical staff in the Performance Engineering team at VMware where she is responsible for characterizing the performance of TPC-E and Exchange on Windows. Prior to VMware, she was a performance engineer at Hewlett-Packard where she worked with OLTP benchmarks such as TPC-C and TPC-DS, and SAP Standard Application benchmarks.

Acknowledgements

The authors would like to thank Bing Tsai, Shilpi Agarwal, Julie Brodeur, Vincent Lin, Priti Mishra, and Todd Muirhead for reviews and contributions to the paper.

