VIRTUALIZING MICROSOFT EXCHANGE SERVER 2007 WITH VMWARE ESX SERVER 3

BY TODD MUIRHEAD
KONG YANG

Before virtualizing Microsoft® Exchange Server 2007, organizations should understand how those virtual machines might perform. This article describes tests used to evaluate the performance and scalability of Microsoft Exchange Server 2007 on a VMware® ESX Server 3 virtualization platform.

Related Categories:

Dell Enterprise Technology Center

Microsoft Exchange Server 2007

Virtualization

VMware

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

Initial enterprise virtualization deployments often focus on test and development and server consolidation. Once these initial deployments are successful, the next step is to identify additional applications that can be virtualized, often beginning with basic, noncritical applications. Finally, enterprises might consider virtualizing their most critical, performance-intensive applications—including Microsoft Exchange, a common application that can be critical to enterprise operations and requires high performance. Evaluating Exchange in virtualized environments can help determine whether other performance-intensive applications are suitable candidates for virtualization.[1]

A previous *Dell Power Solutions* article[2] examined how Microsoft Exchange Server 2003 performed on a VMware ESX Server virtualization platform running on a Dell™ PowerEdge™ 6850 server. Exchange Server 2007, however, introduces many different server roles and features, and as a 64-bit application, it offers higher memory scaling than Exchange Server 2003. To help demonstrate how Exchange Server 2007 performs in a virtualized environment, in August 2007 Dell engineers performed tests designed to address four key aspects of virtual machine (VM) sizing and performance: initial VM sizing, quality of service, appropriate resource levels for different user loads, and whether a single large VM or multiple small VMs provide the highest performance. The results indicated that in the test environment, appropriately configured VMs could support up to 500, 1,000, or 2,000 heavy Exchange users while continuing to provide acceptable quality of service.

Hardware and software configuration in the test environment

To provide an appropriate baseline VM configuration, the test team based the test environment on Dell reference architectures for Exchange Server 2007 and recommendations of the Dell Exchange 2007 Advisor tool.[3] This environment utilized a Dell PowerEdge 2950 server—a 2U, two-socket rack server that supports Intel® Xeon® 5100 and 5300 series processors and up to 32 GB of RAM, and provides a PCI Express (PCIe) riser with three PCIe slots or two PCI Extended (PCI-X) slots and one PCIe slot.

The test team configured this server with two quad-core Intel Xeon E5345 processors at 2.33 GHz, 16 GB of RAM, and VMware ESX Server 3.0.1. Each VM ran the Microsoft Windows Server® 2003 Enterprise x64 Edition OS and Microsoft Exchange Server 2007. All of the Exchange Server 2007 server roles were virtualized.

[1] When considering the deployment of Microsoft Exchange Server 2007 on a virtualization platform such as VMware ESX Server, please refer to support.microsoft.com/kb/897615 for information about the Microsoft support policy for their applications running in non-Microsoft virtualization software.

[2] "Microsoft Exchange Server 2003 Performance on VMware ESX Server 3," by Kong L. Yang and Aravind Pavuluri, in *Dell Power Solutions*, August 2007, DELL.COM/downloads/global/power/ps3q07-20070369-Yang-OE.pdf.

[3] See "Dell Exchange 2007 Advisor and Representative Deployments," by Farrukh Noman and Bharath Vasudevan, Dell Product Group, April 2007, DELL.COM/downloads/global/solutions/Dell_Exchange_2007_Advisor.pdf; and the Dell Exchange 2007 Advisor tool at DELL.COM/content/topics/global.aspx/tools/advisors/exchange_advisor.

The server was connected to a Dell/EMC CX3-80 storage array through McDATA Sphereon 4700 Fibre Channel switches and to client systems through a Dell PowerConnect™ 5324 switch. The client systems included a Microsoft Active Directory® domain controller, a Domain Name System (DNS) server, a VMware VirtualCenter server, and a Microsoft Exchange Load Generator (LoadGen) control server.

The Dell reference architectures for Exchange Server 2007 include small, medium, and large configurations. Figure 1 outlines the basic configurations for each VM. To help evaluate the performance of Exchange Server 2007 in a virtualized environment, the test team ran these small, medium, and large VM configurations in different tests; by examining key metrics during these tests, the team determined the configurations that delivered the highest performance while maintaining good quality of service.

The primary metrics used to compare performance were the host server's processor utilization and the 95th-percentile response time for the LoadGen SendMail task. The test team used Microsoft Windows® Performance Monitor (perfmon) to monitor VM performance and esxtop to monitor host server performance. The team configured perfmon to log processor, memory, disk, network, and Exchange-specific counters, which were logged in a comma-separated values (CSV) performance log file and later imported into a spreadsheet for analysis. The team configured esxtop to log processor, memory, disk, and network counters for both ESX Server and the VMs.

The test team used the LoadGen tool running on a client server to simulate a messaging workload utilizing the included heavy workload profile. LoadGen models the normal daily e-mail usage profile of real users and can help estimate the number of users that a system can support. It reports performance metrics that indicate how an Exchange server would typically respond when supporting a given number of users. LoadGen reports results in terms of response time or latency across 16 common e-mail tasks, including sending messages, deleting messages, browsing a calendar, logging on, and logging off. Although LoadGen tests do not necessarily match real-world deployments, they can provide usable data for comparison between different workloads. In the SendMail task chosen by the test team to evaluate Exchange performance, good quality of service is defined as the 95th percentile's average latency being less than 500 milliseconds.

	Small	Medium	Large
Users	500	1,000	2,000
Virtual processors	1–4	1–4	1–4
Memory	8 GB	8 GB	16 GB
Data disks	4	8	16
Log disks	2	4	8
Virtual network interface cards	1	1	1
Storage groups	4	8	16

Figure 1. *Basic virtual machine configurations in the test environment*

Test results: Response times and processor utilization

By taking advantage of the Dell reference architectures for Exchange Server 2007, the test team determined that the disk, memory, and network resources were sufficient at each configuration size. Therefore, the testing focused on analyzing the effects of virtual processors (vCPUs) on the overall performance of the Exchange VMs. The results across the majority of test cases showed good response times, with the 95th percentile for the SendMail task below or near 500 milliseconds. The response times only greatly exceeded this threshold in configurations where the processors or amount of memory were clearly inadequate.

Scaling up: Increasing virtual processors in one virtual machine

The test results indicated that the optimal number of vCPUs per VM is determined by the number of users that VM needs to support. When the VM was supporting 500 heavy users, one vCPU provided lower response times and more efficient processor utilization than two or four vCPUs (see Figure 2). When the VM was supporting 1,000 heavy users, two vCPUs provided the lowest response times (see Figure 3). When the VM was supporting 2,000 heavy users, four vCPUs provided the lowest response times (see Figure 4). In the 2,000-user tests, one vCPU was insufficient to complete the LoadGen tasks. Overall, the host server's processor utilization remained low—less than 25 percent—across all tests using all three types of VMs.

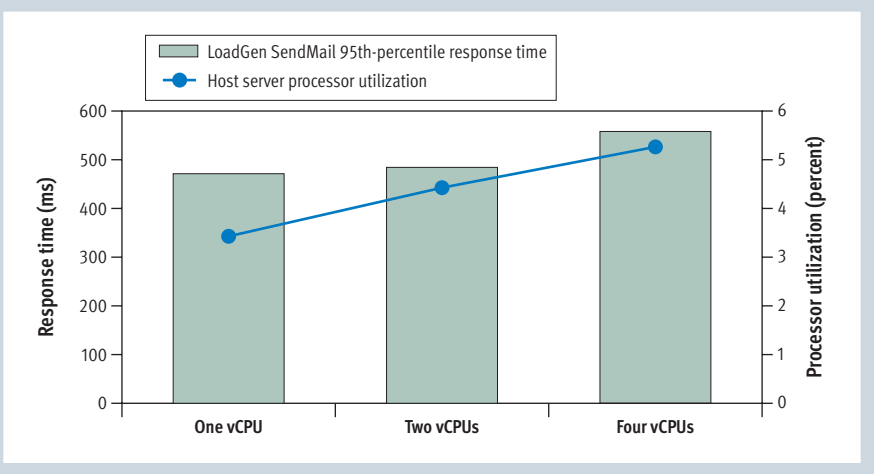


Figure 2. *Scaling up: Different numbers of virtual processors in one small virtual machine supporting 500 heavy LoadGen users*

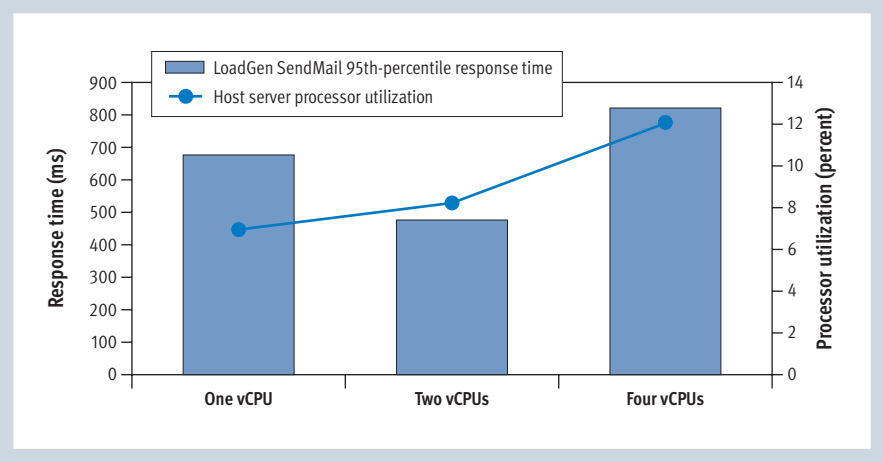


Figure 3. *Scaling up: Different numbers of virtual processors in one medium virtual machine supporting 1,000 heavy LoadGen users*

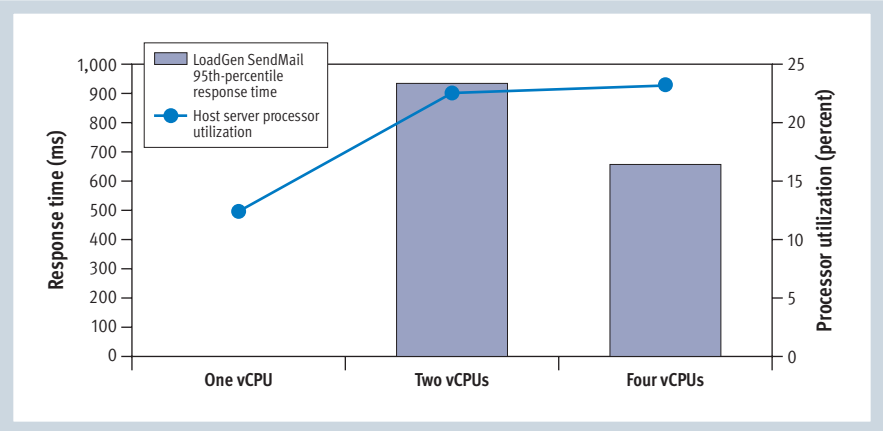
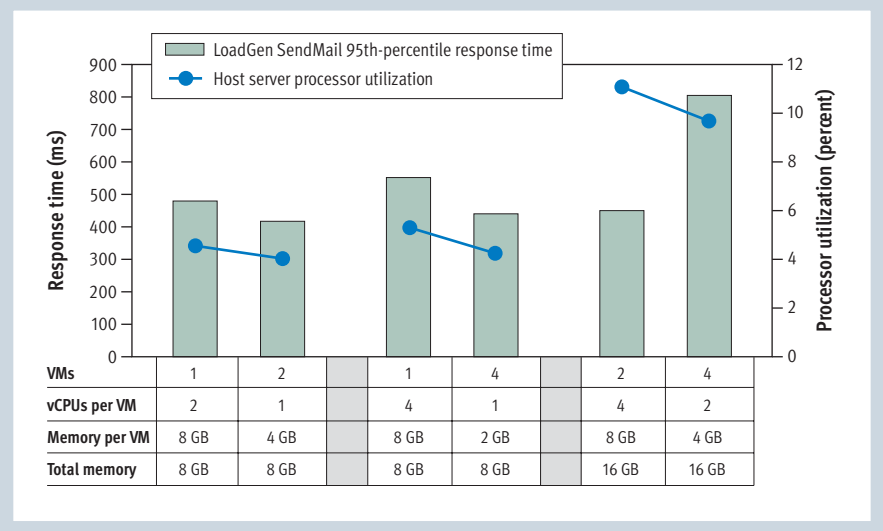


Figure 4. *Scaling up: Different numbers of virtual processors in one large virtual machine supporting 2,000 heavy LoadGen users*



VMs	1	2		1	4		2	4
vCPUs per VM	2	1		4	1		4	2
Memory per VM	8 GB	4 GB		8 GB	2 GB		8 GB	4 GB
Total memory	8 GB	8 GB		8 GB	8 GB		16 GB	16 GB

Figure 5. *Scaling out: Different numbers and configurations of small virtual machines supporting 500 heavy LoadGen users*

Scaling out: Varying virtual machines, virtual processors, and memory

In addition to examining how the VMs scaled up in different configurations with different numbers of vCPUs, the test team also evaluated how these systems scaled out with different combinations of VMs, vCPUs, and memory while keeping the disk and network subsystems constant. Figures 5–7 show the results. Providing sufficient processor and memory resources to the VMs helped reduce overall host server processor utilization and response times. However, assigning processor and memory resources beyond what was necessary tended to degrade performance.

Evaluating qualitative response times

As the preceding results show, although there were a few outliers in cases where the processor resources were insufficient for a particular user load, the tested configurations generally produced acceptable response times, including at least one set of LoadGen SendMail 95th-percentile response times that was near or below the 500-millisecond threshold. To help qualitatively evaluate the difference between a 95th percentile of 500 milliseconds and a 95th percentile of 1 second, the test team opened Microsoft Office Outlook® 2007 client sessions on two terminals, then composed and sent e-mail requests while LoadGen was

TALK BACK

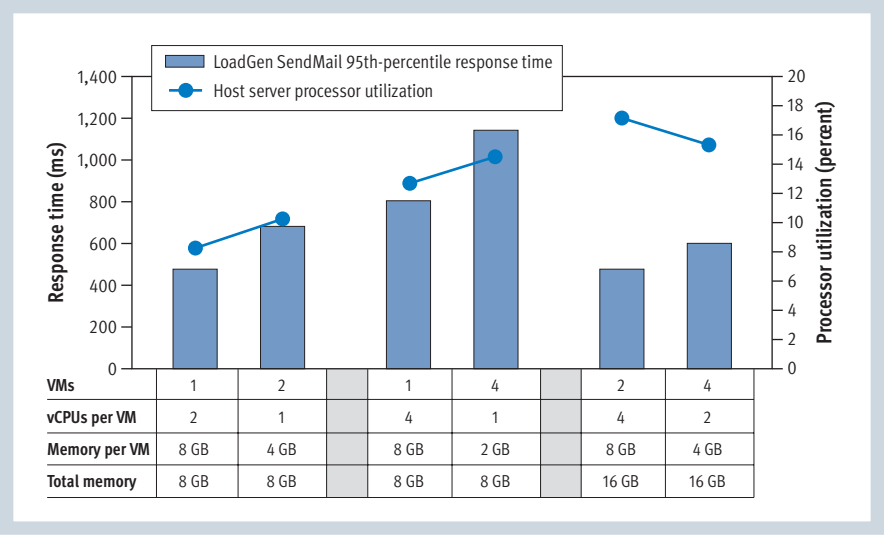
Tell us how the Dell Enterprise Technology Center can help your organization better simplify, utilize, and scale enterprise solutions and platforms. Send your feedback and ideas to enterprise_techcenter@DELL.COM.

running. In each case, as soon as a tester sent a message, the message went to the out-box and was in transit in less than 1 second. A few seconds later, the message appeared in the receiver's in-box. The difference between configurations at a 500-millisecond threshold and those at a 1-second threshold was imperceptible to these testers.

Microsoft Exchange performance in virtualized environments

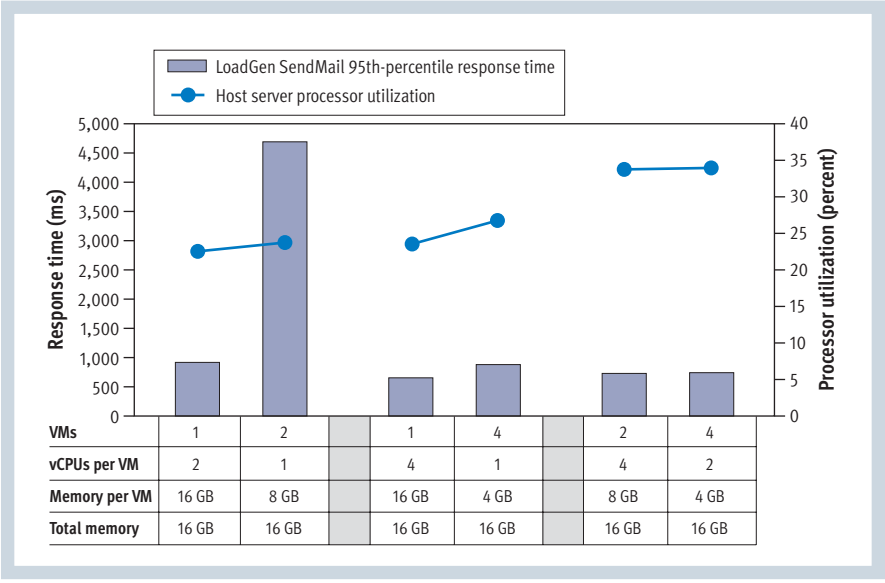
Understanding how Microsoft Exchange, a communication lifeline for many enterprises, performs in virtualized environments is key to planning data center deployments. In the Dell tests described in this article, a VM with one vCPU provided the lowest response times when supporting 500 heavy users, a VM with two vCPUs provided the lowest response times when supporting 1,000 heavy users, and a VM with four vCPUs provided the lowest response times when supporting 2,000 heavy users.

In these tests, the configurations with many small VMs provided performance and response times similar to those with a few large VMs—for this reason, enterprises should base their VM configurations on factors such as software licensing costs, ease of management, and enterprise standards. These factors can be particularly important when utilizing ESX Server features such as VMware VMotion™ and VMware High Availability (VMware HA) technology, because downtime and migration time can affect overall quality of service for end users. Although LoadGen results are not necessarily representative of production deployments, and performance will vary depending on the specific infrastructure, Exchange configuration, and workload profile, these results can provide a reference point for this particular configuration using this specific workload profile.



VMs	1	2		1	4		2	4
vCPUs per VM	2	1		4	1		4	2
Memory per VM	8 GB	4 GB		8 GB	2 GB		8 GB	4 GB
Total memory	8 GB	8 GB		8 GB	8 GB		16 GB	16 GB

Figure 6. *Scaling out: Different numbers and configurations of medium virtual machines supporting 1,000 heavy LoadGen users*



VMs	1	2		1	4		2	4
vCPUs per VM	2	1		4	1		4	2
Memory per VM	16 GB	8 GB		16 GB	4 GB		8 GB	4 GB
Total memory	16 GB	16 GB		16 GB	16 GB		16 GB	16 GB

Figure 7. *Scaling out: Different numbers and configurations of large virtual machines supporting 2,000 heavy LoadGen users*

Todd Muirhead *is a senior engineering consultant at the Dell Enterprise Technology Center.*

Kong Yang *is a systems engineer in the Dell Virtualization Solutions Engineering Group.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Dell Enterprise Technology Center:
DELL.COM/TechCenter

Dell Enterprise Technology Center wiki:
www.delltechcenter.com

Dell and VMware:
DELL.COM/VMware

Microsoft Exchange solutions on Dell:
DELL.COM/Exchange