VeriTest®

January 2009

# Sun Microsystems Sun Fire X4600 M2 server: Performance and Scalability for Virtual Desktop Infrastructure

***Independent Test Report Prepared for Sun Microsystems***

## Executive Summary

Sun Microsystems, Inc. (hereinafter "Sun") commissioned VeriTest, the testing service of Lionbridge Technologies, Inc., to conduct a study aimed at demonstrating a theoretical maximum for the number of "Desktop" type virtual machines (VMs) commonly used in a Virtual Desktop Infrastructure (VDI) that can be supported on a single Sun Fire X4600 M2 server of a specific configuration running on VMware Infrastructure 3 (ESX 3.5.0 Build 110181 with VirtualCenter Server 2.5). Testing was performed at the Lionbridge test facility in Oakdale, Minnesota.

### Key Findings

- Testing showed that the Sun Fire X4600 M2 server was capable of supporting as many as 160 Windows XP Desktop Virtual Machines when under typical desktop load.
- Testing showed that at 160 desktop virtual machines the overall CPU utilization was 93%, allowing for additional virtual machines before reaching the 95% target utilization rate.
- The Sun Fire X4600 M2 server encountered no alerts of any kind from either ESX or the system when running the 140 Virtual Machines. CPU alerts began to appear for brief times with 160 VMs.
- Due to the nature of the WorldBench benchmark being geared to run its specific workload at a higher rate than a real-world user and the fact that at 160 VMs the CPU utilization was at 93%, it is believed that more than 160 VMs could be supported.

Virtual Desktop Infrastructure, commonly known as VDI, is a computing model that is designed around the architecture of centralizing the desktop computing resources into the data center and replacing each desktop with a more secure and cost-efficient device. The basic concept is to shift physical desktop workstations and their corresponding workloads at each user desk space, to independent and dedicated virtual machines that are residing on secure, robust servers in a centralized data center. Each specific end-users desktop operating environment is then remotely accessed from a more secure and more easily manageable device, such as a thin client. This model is intended to ensure data security, increase the administrators ability to better manage and service each individual end-user, all while providing secure access to the end-users identical environment they are already familiar with, such as Windows XP or Vista.

Testing was performed at different levels of total load with a goal of achieving as many VMs on a single Sun Fire X4600 M2 server as possible without exceeding an average CPU utilization of over 95% for the duration of the test simulation. In our tests, this resulted in a simulation of 120 VMs, 140 VMs, and finally 160 VMs. In an effort to appropriately simulate real-world usage, testing was focused on workloads that are most commonly seen in an average VDI desktop user scenario.

PCWorld's WorldBench 5 tool was used to generate typical Windows user load on the desktop VMs. WorldBench 5 is the fifth generation of PC World's industry-standard benchmarking application designed to measure the performance of a wide range of personal computers. WorldBench 5 uses real applications running real-world tasks to assess a PC's overall performance.

## Testing Methodology

VeriTest developed a test methodology in which Windows XP x86 (SP3) Virtual Machines were configured with a workload that simulated an average desktop user experience. To simulate this workload, VeriTest used the PC World WorldBench v5 application. The WorldBench benchmark is able to simulate various desktop workloads such as office applications, multimedia and internet activity. Since testing was focused on typical enterprise desktop environments, the Office XP workload simulation was selected as most appropriate. This workload simulates the usage of the core Office suite applications Word, Excel, PowerPoint, Outlook and Access. These applications are each launched and then perform common desktop tasks such as opening, closing, creating, and modifying .doc, .xls and .ppt files as well as composing e-mails. During the workload, all Office applications are opened simultaneously and a variety of tasks are executed. These include such things as:

- Reading, composing, deleting, and sending mail, creating and viewing calendar appointments, entering, modifying, and deleting contacts, creating and editing tasks in Microsoft Outlook.
- Typing and formatting text, scrolling from page to page, running spell check, conducting mail merges, print previewing documents in Microsoft Word.
- Performing calculations, creating charts, sorting data, formatting the screen, and previewing spreadsheets in Microsoft Excel
- Creating slides, formatting slides, modifying the master template, adding, editing, and deleting text, and viewing slideshows in Microsoft PowerPoint
- Entering data, executing queries, and generating reports in Microsoft Access

It should be noted that WorldBench is designed to run the various simulations simultaneously at the fastest rate the target system can execute the program. This rate of execution is indisputably a faster rate than a "real" user would execute the same various applications and tasks.

Each WorldBench instance was configured to execute the Office-XP profile for a total of 15 cycles. The number of cycles was used to ensure that 3 hours of test execution time could be captured to best determine the maximum number of VMs that could be supported on the Sun Fire X4600 M2 server with the simulated workload.

Each World Bench cycle contains numerous system and application setup stages prior to, and after, the execution of the benchmark suite. These tasks are performed on each VM running the WorldBench benchmark suite for Office XP. Each of these stages are executed by the WorldBench application as part of the test cycle. These WorldBench stages include the following:

1. Backup the registry
2. System Modification by WorldBench
3. Reboot
4. Setup system files
5. Setup alluser files
6. Setup user files
7. Reboot
8. Registry setup for WB
9. Setup application files -> msoffice
10. Setup registry for msoffice
11. Clear system buffers
12. Clear background tasks
13. Wait for system to become idle
14. Execute benchmark
15. Remove application files
16. Setup alluser files
17. Setup user files
18. Setup the registry for WB
19. Setup application files -> msoffice
20. Setup registry for msoffice
21. Clear system buffers

22. Clear background tasks
23. Wait for system to become idle
24. Execute benchmark
25. Steps 15-24 repeated for each test cycle
26. Remove application files
27. Restore alluser files
28. Restore user files
29. Restore system files
30. Clean up the registry
31. Reboot
32. Test complete

WorldBench 5 combines the results of scripted application tests and then compares them to the scores of a reference system – published as a system with a 2.2-GHz Athlon 64 FX-51 CPU with 1MB of Level 2 cache and 1GB of RAM, as well as an NVidia GeForce FX 5950 Ultra graphics card with 256MB of RAM. For the purposes of this study, WorldBench was used as a load generation tool to simulate workload that is most common in an enterprise desktop environment. Due to the rapid nature of the application tasking WorldBench performs, the load placed on the ESX host by the VMs simulation was expected to be greater than what a typical user would achieve using the same applications. Such a situation allows for some favorable conclusions to be made with regard to how many VMs can be supported.

WorldBench is a tool designed for the physical environment and cannot take into account the nuances of shared resources in a virtual environment. As a result, the overall WorldBench score was not relevant to the testing. The focus of this study was on the performance of the server under test. However, during all test runs, the individual VMs were sampled to ensure their application performance fell within a reasonable range.

In addition, since WorldBench runs locally on the target system, the effects associated with accessing the Windows XP VMs over a network connection (such as from a Sun Ray virtual display client) was not considered as part of the performance associated with the Sun Fire X4600 M2 platform under test.

## *Test Bed*

Figure 1 below shows the test environment configured in support of this testing.

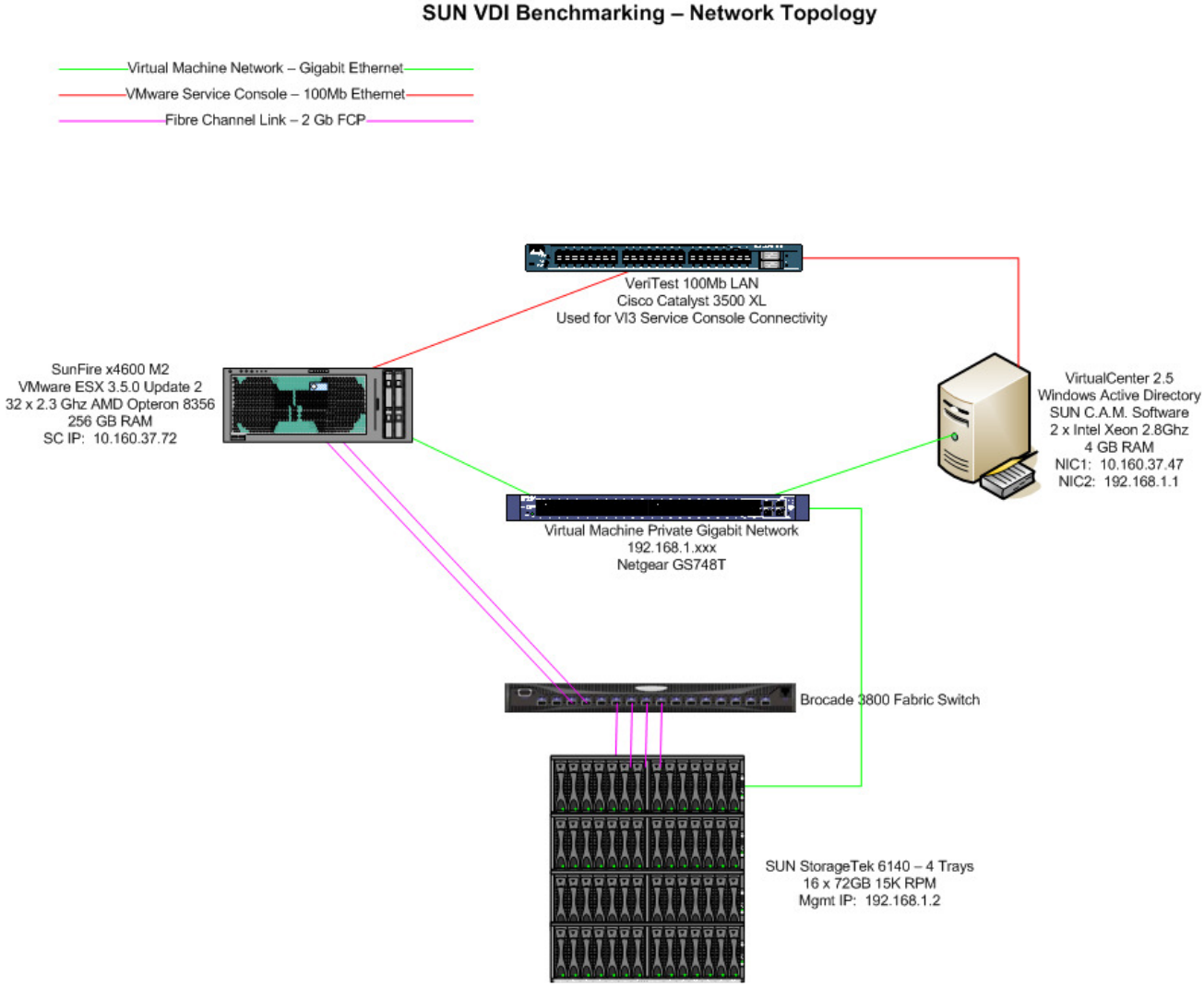


**Figure 1: Test Bed Diagram**

The server under test was a Sun Fire X4600 M2 server. This system contained eight (8) Quad-core AMD Opteron 8356 processors running at 2.3 GHz, 256 GB RAM, 4 GbE Ethernet ports, and 4 x 72 GB 10K RPM SAS drives. VMware ESX Server 3.5.0 Update 2 was installed onto one of the internal disks and a default Datastore configured. This local Datastore was configured to hold the template Windows XP x86 w/SP3 Virtual Machine image.

To provide storage for the Virtual Machines, a Sun StorageTek 6140 storage array was configured with 8 volumes using the Microsoft_NTFS profile. This configuration allocated 4 disks in a 3+1 configuration using RAID-5 for each of the 8 volumes. A total of 20 Virtual Machines could be allocated to each volume if needed. The STK6140 was in a dual controller configuration. Each controller had two Fibre Channel ports connected to the Fibre Channel switch. In addition, the Sun Fire X4600 M2 server had a single dual-port QLA2342 HBA installed. Each port of the HBA was connected to the Fibre Channel switch. The Fibre Channel switch was configured with two zones that dedicated port1 from the server's FC HBA to port1 on the STK6140 Controller A and port2 of the server's FC HBA to Controller B port1 on the STK6140. This resulted

in a configuration in which the 8 storage volumes were evenly balanced across the two FC HBA ports on the Sun Fire X4600 M2 as well as evenly balanced across the two controllers of the STK6140.

To represent the virtualized desktops, VeriTest first installed a single source Virtual Machine (VM) which contained the Windows XP x86 with SP3 Operating System along with the WorldBench 5 application. Each VM was configured with a single vCPU, 512 MB of RAM, a single 8 GB HDD drive image, and a single NIC using the Flexible adapter. This image was then replicated using the VMware VI3 cloning capabilities. Each cloned Virtual Machine image was then made unique by modifying the system hostname and then joining the VM to the Windows Active Directory domain.

A separate physical server was used to run Windows Active Directory services. This was equipped with dual 2.8 Ghz Intel Xeon CPUs, 4 GB of RAM, dual Gigabit Ethernet NICs, and 2 x 36 GB 15K RPM U320 SCSI HDDs in a RAID-1 configuration running Windows Server 2003 R2 Enterprise Edition with SP2 x86. This system was used to provide Active Directory services to the Windows XP Virtual Machines. This system also had VirtualCenter Server 2.5 Update 2 installed.

Finally, to measure the CPU and RAM performance of the Sun Fire X4600 M2 server the VMware VI3 utility esxtop was utilized. The command was run in batch mode to capture all performance metrics every 10 seconds. In addition to the esxtop command running, the Sun Fire X4600 M2 server was monitored with VirtualCenter to ensure that no warnings or alerts were reported by ESX with regard to other monitored resources such as RAM usage.

## *Test Execution*

Based upon CPU resources and the overall X4600 M2 system configuration, it was determined to start the testing using 120 Virtual Machines and ramp up from there. The goal was to load the system up with as many virtual machines as possible while not exceeding a maximum CPU load of an average of 95%. This allows for some overhead room to exist in the event a small number of the VMs experience a short spike in usage.

120 Virtual Machines were initiated with the WorldBench workload simulation. During this process, an esxtop session was run in batch mode to capture the various metrics reported. In addition, a real-time esxtop session was run to allow the VeriTest Test Engineer to monitor the CPU and RAM load placed on the Sun Fire X4600 M2 server. Upon completion of the 120 VM test, a second test was executed with 140 VMs, and a third test was executed with 160 VMs. All tests were executed twice to ensure repeatability of the results. The load was allowed to run for a minimum of three hours for each test run to ensure the average CPU load was sustained as close to or at the 95% range for this extended period of time.

In addition to CPU load, the Sun Fire X4600 M2 server was monitored to ensure there were not any other performance related alerts.

## Test Results

As described in the Test Methodology section above and illustrated in Figure 2 below, the first test performed involved running WorldBench on 120 virtual machines. In this test scenario, the WorldBench simulation was executed over approximately a three hour timeframe. The average CPU utilization was collected from VMWare's esxtop and was sampled every 10 seconds. The Y axis on the chart shows the execution time in "hours:minutes:seconds" beginning at zero as the start of testing and continuing to the final sampling time roughly three hours later.

In order to produce a test cycle that would run over a three hour period, WorldBench was configured to run the OfficeXP workload fifteen times. At the beginning of each test cycle, WorldBench performs a variety of configuration changes and updates and then reboots each VM. These changes are subsequently reset to the original settings at the end of the test cycle and each VM is again rebooted. First-hand observations by VeriTest engineers indicate the "valleys" in Figure 2 below appear to be directly correlated to the WorldBench application preparation stages, steps 1 thru 13, as described in the Test Methodology section above. The peaks in the below figure are the periods in which WorldBench was in the process of executing the Office XP workload simulation.

Testing showed that when running 120 virtual machines, the average CPU utilization was roughly 74%; thus allowing the addition of more VMs.

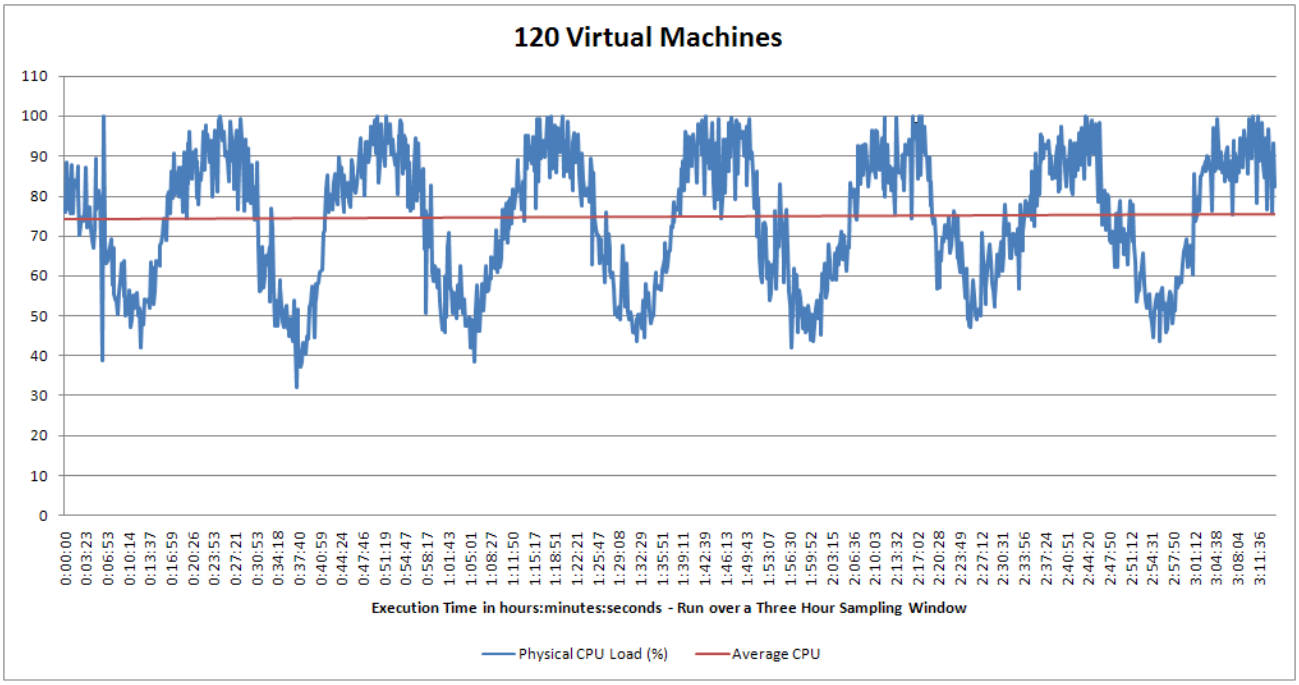


**Figure 2: CPU Load - 120 Virtual Machines**

Next, a test of 140 Virtual Machines was executed in an attempt to isolate the maximum number of virtual machines the Sun Fire X4600 M2 server might support. In this test scenario, the WorldBench simulation was once again executed over approximately a three hour timeframe. The CPU utilization was collected from VMware's esxtop and was sampled approximately every 10 seconds. The Y axis on the chart shows the execution time in "hours:minutes:seconds" beginning at zero as the start of testing and continuing to the final sampling time roughly three hours later.

As seen in Figure 3 below, when running WorldBench with 140 VMs, the CPU utilization was higher than observed in the 120 VM test and averaged 79%. As a result, a final test cycle was executed in order to push the X4600 M2 further.

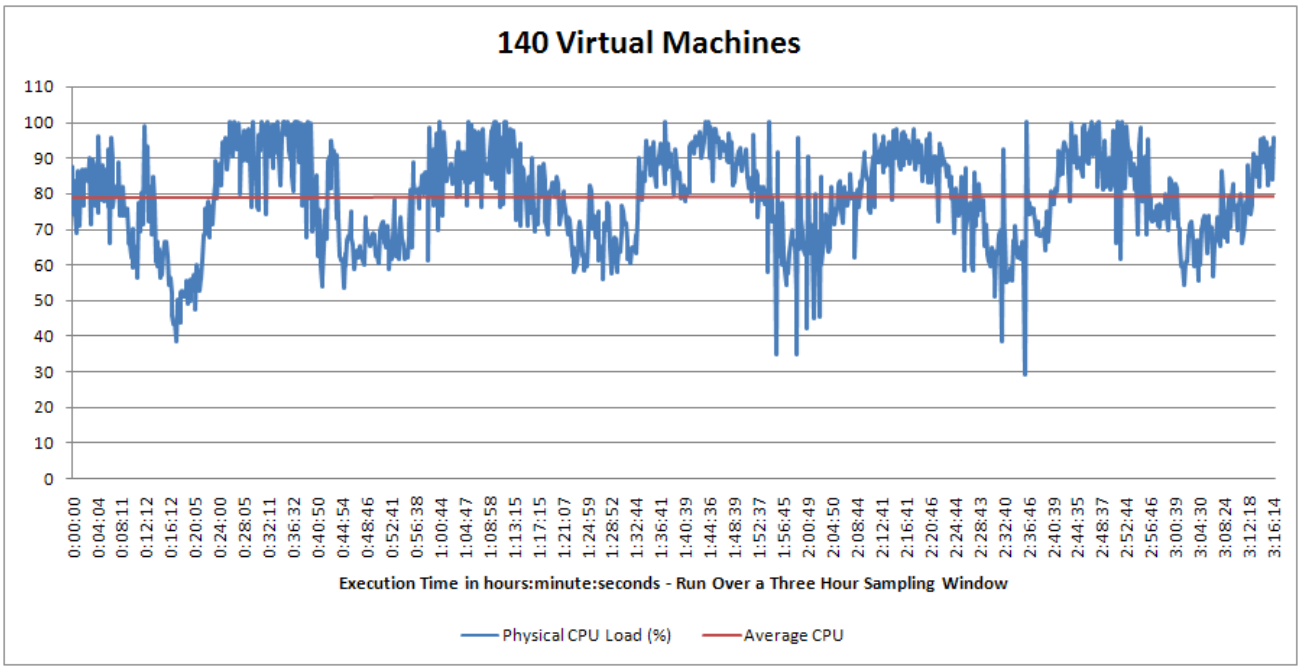


**Figure 3: CPU Load - 140 Virtual Machines**

One final test was executed to attempt to reach the 95% CPU utilization mark on a single X4600 M2. In this test scenario, the WorldBench simulation was executed over approximately a three hour timeframe. The CPU utilization was collected from VMware's esxtop and was sampled approximately every 10 seconds. The Y axis on the chart shows the execution time in "hours:minutes:seconds" beginning at zero as the start of testing and continuing to the final sampling time roughly three hours later.

As illustrated in Figure 4 below, when running 160 virtual machines during the test cycle it was noted that the CPU utilization was not measurably higher than when running the 140 VMs. In fact, the average CPU utilization was 80%. Additionally, it was noticed that when sampling individual desktop virtual machines that their responsiveness had degraded and was below what had been observed during the 140 VM test. This can be seen in the broader declines in the chart showing those times the virtual machines were rebooting and the increased length in which it took an individual VM to get back to the point of executing tests.

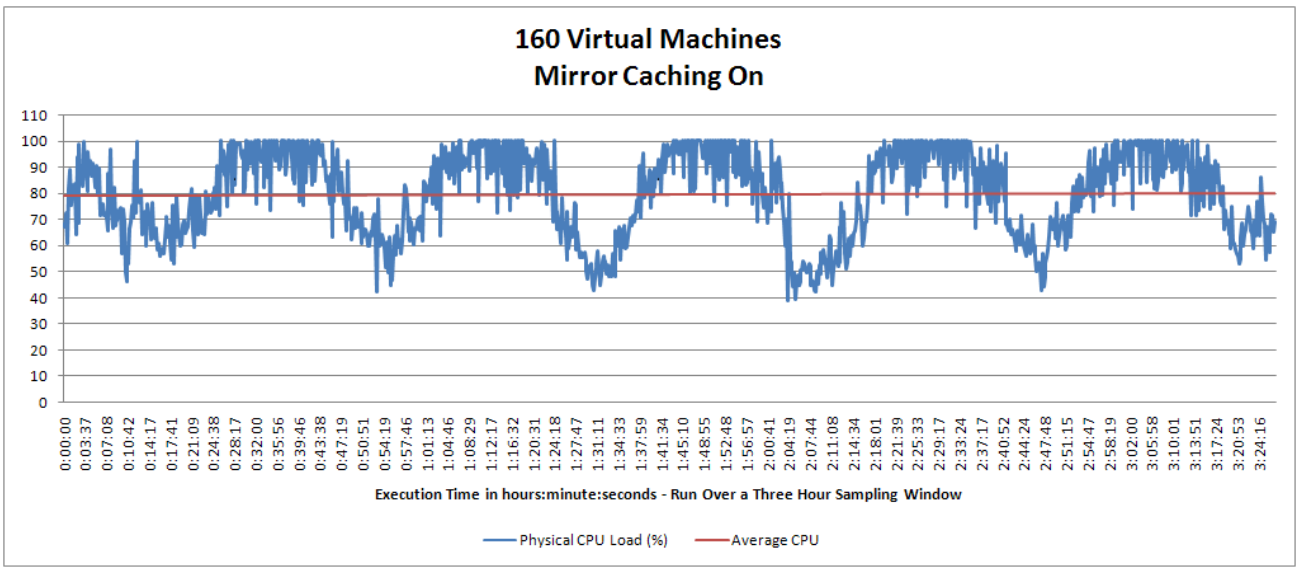


**Figure 4: CPU Load - 160 Virtual Machines**

Upon completion of this test cycle, it was determined that a bottleneck existed elsewhere in the environment. Upon investigation, it appeared that the particular storage array configuration in place for this test was at the root of the performance. Through conversations with SUN technical representatives, it was suggested that in this benchmark configuration performance may benefit from the disabling of the storage controller Mirror Caching. Mirror Caching provides an additional layer of data protection in the environment that is generally required for production environments, but is not necessarily required in benchmark situations where the storage array is not the product being benchmarked.

As you can see in Figure 5 below, this additional test validated the suspicion that there was a bottleneck in the environment. Disabling the Mirror Caching delivered the additional throughput on the storage that was needed to allow the individual desktop virtual machines to perform appropriately.

This test demonstrated an overall average load of 93%. Additionally, the responsiveness of the individual virtual machines was restored to what had been previously observed in the 120 and 140 virtual machine tests.

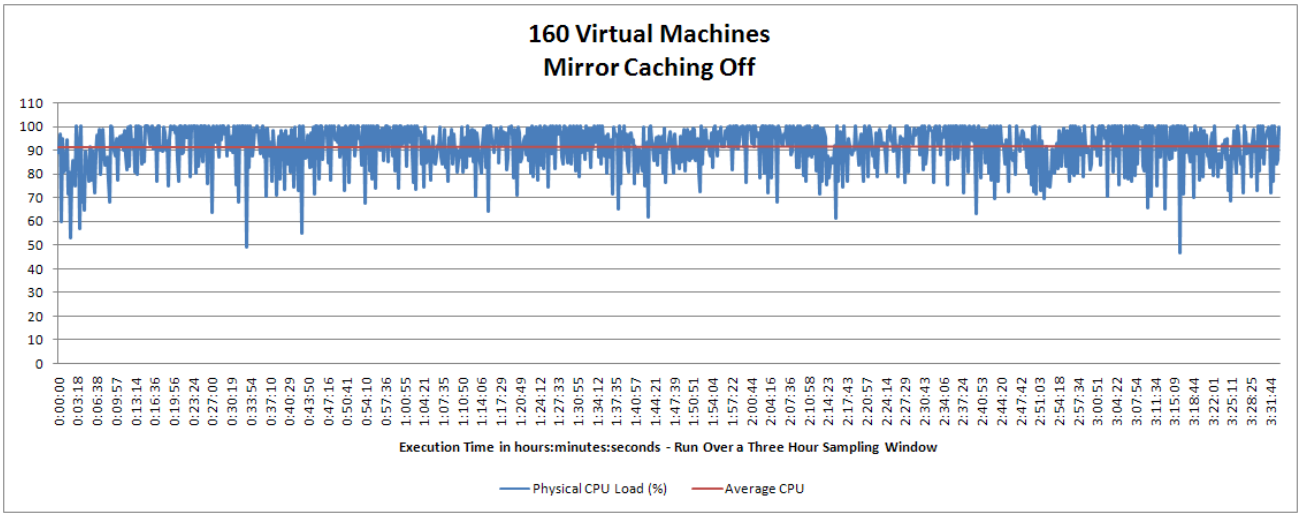


**Figure 5: CPU Load - 160 Virtual Machines - No Mirror Caching**

Although this test did not push the X4600 M2 CPU to the target utilization rate of 95%, due to capacity limitations within the specific test environment that was configured, 160 VMs represented the maximum number of virtual machines that could be loaded. However, the testing demonstrated that the X4600 M2 can comfortably support 160 VMs, and given the workload WorldBench administers, could likely support close to the currently maximum number of VMs supported by ESX 3.5 of 192.

It should be noted that during all test simulations, no ESX alerts were recorded pertaining to resource contentions with the Sun Fire X4600 M2 server. ESX was configured to issue an alert whenever the ESX host would experience CPU load that reach 95% or above. The only alerts that appeared were with regard to CPU utilization, which was expected when the load began to reach the 95% plus mark during the second running of 160 VMs.

## Conclusion

Testing showed that in the environment under test, the Sun Fire X4600 M2 server was capable of supporting as many as 160 Virtual Machines running the WorldBench Office XP workload simulation, at an average CPU utilization of 93%. This represented the maximum number of virtual machines that could be loaded in the current test environment. With more storage resources, it is reasonable to conclude that additional virtual machines could be loaded on this same X4600 M2.

Since the WorldBench simulation generates a more consistent and intensive workload than might typically be found in the enterprise, it is reasonable to expect customers could achieve an even greater number of desktop virtual machines on a single Sun Fire X4600 M2 server in their environment.

# Appendix A. Hardware Disclosures

### *Server Under Test*

| | |
|---|---|
| **Model** | Sun Fire X4600 M2 server |
| **BIOS** | 0ABIT087 Date: 03/19/08 Core: 08.00.12 |
| **CPU** | 32 x AMD Opteron 8356 2.3 Ghz |
| **RAM** | 256 GB, DDR2 667 Mhz |
| **HDD** | 4 x 72 GB 10K RPM SAS |
| **NIC** | 4 x 10/100/1000 Onboard |
| **Fibre Channel HBA** | QLA2342 |
| **BIOS** | 1.04 |
| **ESX Version** | 3.5.0 Update 2 Build 110181 |

### *Storage Array*

| | |
|---|---|
| **Model:** | Sun STK6140 |
| **Firmware:** | 6.60.11.10 |
| **Controllers:** | 2 |
| **Trays:** | 4 |
| **Disks:** | 64 x 72GB 15K RPM FC-SCSI |
| **Volumes:** | 8 |
| **RAID Type:** | 5 |
| **RAID Profile:** | Microsoft_NTFS |
| **Number of Disks** | 4 (3+1) |

### *Fabric Switch*

| | |
|---|---|
| **Model** | Brocade 3800 Silkworm |
| **Firmware** | 3.2.0a |

### *Network Switch*

| | |
|---|---|
| **Model** | Netgear GS748T |
| **Firmware** | V2.0.2_02 |

### *Windows Domain Controller*

| | |
|---|---|
| **CPU** | 2 x 3.0Ghz Intel Xeon |
| **RAM** | 4 GB |
| **HDD** | 2 x 72GB 15K RPM U320 SCSI (RAID-1) |
| **NIC** | 2 x 10/100/1000 Broadcom (onboard) |
| **NIC1** | Remote Admin Access |
| **NIC2** | VM Network Access |
| **OS** | Windows Server 2003 EE SP2 x86 |

### *Virtual Machine Configuration*

| | |
|---|---|
| **OS** | Windows XP Pro SP3 x86 |
| **vCPU** | 1 |
| **RAM** | 512 MB |
| **HDD** | 8 GB (Bus Logic) |
| **NIC** | 1 (Flexible Adapter) |
| **Benchmark Software** | PC World WorldBench v5.0 |

VeriTest (www.veritest.com), the testing service of Lionbridge Technologies, Inc., provides outsourced testing solutions that maximize revenue and reduce costs for our clients. For companies who use high-tech products as well as those who produce them, smoothly functioning technology is essential to business success. VeriTest helps our clients identify and correct technology problems in their products and in their line of business applications by providing the widest range of testing services available.
VeriTest created the suite of industry-standard benchmark software that includes WebBench, NetBench, Winstone, and WinBench. We've distributed over 20 million copies of these tools, which are in use at every one of the 2001 Fortune 100 companies. Our Internet BenchMark service provides the definitive ratings for Internet Service Providers in the US, Canada, and the UK.

Under our former names of ZD Labs and eTesting Labs, and as part of VeriTest since July of 2002, we have delivered rigorous, objective, independent testing and analysis for over a decade. With the most knowledgeable staff in the business, testing facilities around the world, and almost 1,600 dedicated network PCs, VeriTest offers our clients the expertise and equipment necessary to meet all their testing needs.
**For more information** email us at info@veritest.com or call us at 919-380-2800.