Performance Characterization of the Dell Flexible Computing On-Demand Desktop Streaming Solution

Product Group
Client Business Solutions

Dell White Paper

November 2007

Introduction

Businesses face increasing challenges in deploying, managing and securing their end user computing machines. Today's information workers rely heavily on their corporate desktops or notebooks to accomplish day-to-day tasks and functions. Managing end user systems is a crucial part of business IT operations. The Dell™ Flexible Computing Solution for On-Demand Desktop Streaming, based on the Citrix Provisioning Server for Desktops, enables diskless client computers to boot from a network server via well known mechanisms. The network server, referred to as the "streaming server," delivers disk images containing operating system and applications on demand to clients over the network.

The solution removes the traditional requirements of deploying and maintaining software on a client machine's local hard drive. This helps improve security, manageability and reliability aspects of IT operations for end-user machines in a cost effective manner. Dell has conducted performance testing and characterization of the Flexible Computing Solution for On-Demand Desktop Streaming on supported Dell hardware. This whitepaper highlights some of the observations made, as well as collective knowledge gained through these laboratory experiments, and provides general guidelines for solution resource planning.

Solution Components

The Dell Flexible Computing Solution for On-Demand Desktop Streaming includes streaming servers running Citrix Server software, client machines , the infrastructure components such as Active Directory / DHCP/DNS and network switch components. The various solution components have performance implications on the overall operation of the solution which will be outlined and detailed later in this paper.

Appropriate planning and tuning of these components is necessary to ensure optimal functional performance of the solution. Further, business IT operations implement best practices that require the overall performance of the solution and that of the individual components to be within certain limits. The following sections describe the operational implications of the solution components.

Server

The streaming servers maintain and deliver client hard disk virtual disk images called "vDisks". A vDisk image is prepared from a "master client's" hard drive and can be reused across identically configured client machines. A master client is a machine identical to the clients to be deployed with the exception of a hard drive being present. It must be identical to the clients to be deployed to ensure system driver compatibility. This machine should be configured with the operating system, software and settings desired for the clients to be deployed.

The server can be configured to deliver images in private mode or in standard mode. Private mode delivers one streamed image to one client. Operating in this mode increases the physical security of the client's data. Standard mode delivers one streamed image to more than one client. In addition to the physical security gains, users implementing standard mode also decreases the number of images to maintain. Updates to a single vDisk can be applied to multiple clients when the clients reboot. For the balance of this document, all of the validation and characterization of the solution was performed in standard mode.

Client disk read and write requests are redirected to the streaming server via the network. The server satisfies the read requests by delivering the required information from the read-only vDisk image and satisfies the write requests by writing data to a temporary write cache located on a server disk volume. The contents of the write cache will be deleted after a client shuts down.

The server incurs a processing load during the initial boot of client machines and also during normal operations. The boot processing load is caused by delivery of the initial OS image to the client. The processing required increases with the number of clients booting simultaneously, resulting in increased server utilization levels.

Similarly, during normal client operations the server utilization increases with both the number of clients performing operations and the number of operations performed by those clients. Appropriate capacity planning is required for server resources such as processors, memory, disk and network to ensure the utilization levels and response times meet expectations. High server utilization levels may affect the response times to client operations as well.

Clients

The vDisk image prepared for clients includes the client streaming software. The Citrix client software must be installed on the master client before creating the vDisk image. This allows the creation of a "gold vDisk image" that is made from the master client. That image is then distributed to each identically configured client, which enables clients to communicate with the streaming server for all read and write operations. It redirects all local system drive (usually drive C) read and write requests to the streaming server as network calls. Figure 1 below illustrates this concept. In the "standard-image" mode, disk writes to the local system drive are redirected, kept on a temporary write cache location and deleted at the end of the user's session. The local system disk (if present) can be used for writing temporary data. Persistent application data should be stored on appropriate network shares or locations outside of the streamed image. During operation, clients utilize local resources such as CPU, memory and GPU for all computational operations and processing. Disk requests are the only local operation redirected to the streaming server.

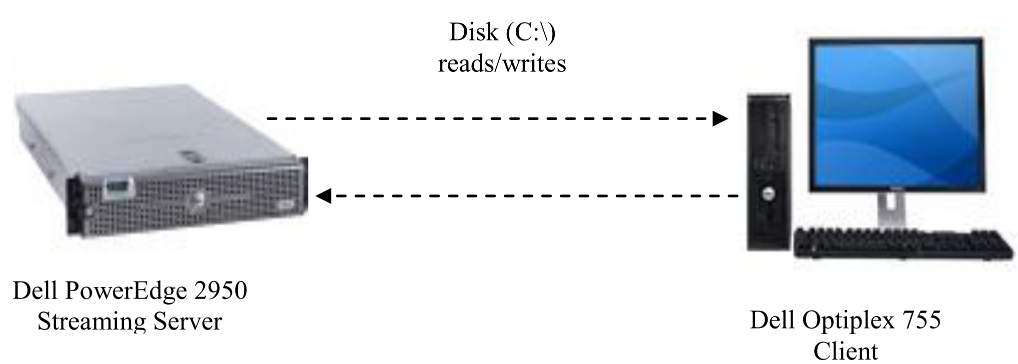

Figure 1: Client disk requests to streaming server

Typically, during boot, the client requests the initial OS binaries required for booting and displaying the desktop. Additional OS binaries, application binaries and data are requested by the client on demand as required during operations. Because of this, during boot and normal operations, clients generate load on the server and affect server response times. This load is proportionate to the number and operations of clients. The load on the server and resultant response times may affect the user experience on each client in terms of time taken to boot the client and retries performed by the client during operations. This data will be covered in more detail in the results section of this paper.

Appropriate server resource planning is required to help ensure that the client experience is not significantly affected due to disk requests being redirected to the streaming server. Resource planning also differs from other flexible computing solutions given that the clients use of local memory, processing, and graphics.

Network Infrastructure

During boot and normal operations, the clients exercise the network infrastructure in addition to exercising server resources. The client disk requests are redirected as network calls on the server, which generate network load in addition to normal client network requests required for applications. Appropriate planning of the network infrastructure is required to handle this additional client load on the network. In our validation, we characterized both an end-to-end Gigabit Ethernet infrastructure as well as an environment with fast Ethernet to the Optiplex Clients with a Gigabit uplink to the datacenter. All of the components validated as part of the On Demand Desktop Streaming solution have embedded Gigabit Ethernet interfaces.

Test Methodology

To illustrate and characterize the concepts noted in the previous section, performance tests were conducted in a highly controlled lab environment with objectives to minimize bottlenecks and to measure performance of known components.

Test Infrastructure

The test infrastructure for the performance characterization studies is represented in Figure 2 below.

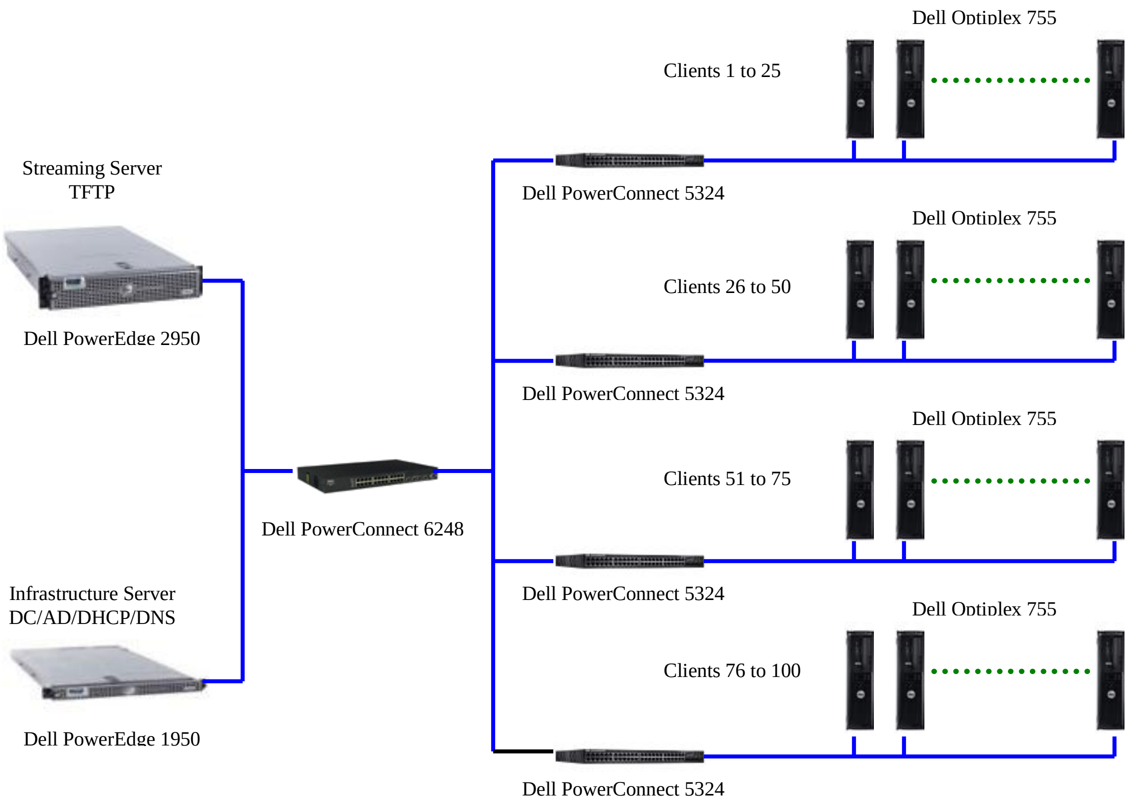

Figure 2: Test Infrastructure

The configuration details of the test infrastructure components are given below:

IO Server (second streaming server used for HA only)

- Dell™ PowerEdge™ 2950
- 1 x Intel® Xeon® 5160 (Dual Core 3.0 GHz)
- 4 GB RAM
- Broadcom dual onboard gigabit network controller

- 6 x 300 GB 10K RPM SAS drives (only 2 required for HA)
    - 1 x 2 disk RAID 1 volume for OS and applications
    - 1 x 4 disk RAID 10 volume for images and cache space (standard mode only)
- Microsoft® Windows Server® 2003 R2 Standard Edition (32-bit)
- Ardence Profession Edition Server 4.1 SP2 (to be renamed Citrix Provisioning Server for Desktops)
- "Standard Image" vDisk mode with one shared image streamed to all 100 clients

Storage Server (HA only)
- Dell PowerEdge 2900
- 1 x Intel® Xeon® 5160 (Dual Core 3.0 GHz)
- 4 GB RAM
- Windows Storage Server 2003 Standard Edition (64-bit)
- Broadcom dual onboard gigabit network controller
- 10 x 146GB 10K RPM SAS drives
    - 1 x 2 disk RAID 1 volume for OS applications
    - 1 x 8 disk RAID 10 volume for vDisk image files, streaming client write cache files, and the streaming server database

Infrastructure Server (can be repurposed from an existing infrastructure)
- Dell PowerEdge 1950
- Windows Server 2003 R2 Standard Edition (32-bit)
- Configured as Domain controller/Active Directory with Microsoft DHCP and DNS services

Client Systems:
- Dell Optiplex™ 755
- 1 x Intel Pentium® D or Core2 Duo (not Celeron®)
- Varying quantities between 512 MB – 2 GB RAM
- Broadcom onboard gigabit network controller
- Ardence Professional Edition Client 4.1 SP2

Primary Switch
- Dell PowerConnect™ 6248
- Spanning Tree on with PortFast enabled

Secondary Switches
- Dell PowerConnect 5324
- Spanning Tree on with PortFast enabled

Test Process

All characterization was performed on the setup described in the "Test Infrastructure" section above. Test cases were designed to test and measure two major categories using multiple scenarios - client experience and server resource utilization. Client experience was characterized by tracking two metrics provided by the streaming software: boot time and session retries. Boot time is a value that signifies the time elapsed from the client's first network contact with the server until the client agent starts on that system once booted into Windows. Session retries is a number which represents how many times the client had to resend a UDP packet to the server because it did not receive a response before the initial packet timed out. Session retries tend to occur when the streaming server and/or the network is busy, which delays response to client requests. The session retries are handled by the client software in the background and can be used to characterize system performance.

Test Scenario 1
The first test scenario in which the two categories were measured is client boot characterization. This involved executing a "boot storm" created by booting varying numbers of clients simultaneously via wake on LAN from an ACPI S5 (off) power state. The purpose of this test was to characterize the behavior of the infrastructure during a peak load of multiple clients booting concurrently and its effect on the client end user experience. Clients were booted in groups of 1, 5, 10, 25, 50, 75 and 100 to establish a baseline and trend for the relationship between group size booting and impact on infrastructure.

Test Scenario 2
The second test scenario is client operational testing, and was performed by booting all 100 clients, allowing the "boot storm" phase to pass, dividing the clients into groups, and then performing different operations on those individual groups. The operations for the client groups included copying files with various source and destination locations for certain groups and rebooting other groups. This test is intended to simulate a heavy load in a "worst case" real world environment which includes multiple clients performing high levels of reboots, application execution and file downloads.

The clients were divided into the following groups during testing

- 25 clients copying while 25 reboot
- 25 clients copying while 50 reboot
- 50 clients copying while 25 reboot
- 50 clients copying while 50 reboot

The copy operations were performed using combinations of local disk and remote network locations as source and destination. Local disk copy included file transfers to or from the client local hard drive which stresses the vDisk and write cache residing on server. This test simulates local drive access in a traditional client system, which typically is exercised for application execution and operating system operations. Remote network location copy operation simulates client access to network shares, mapped drives or any other source or destination for network traffic that is not on the local client machine

The copy operations included clients copying files from:

- Local vDisk to local vDisk to stress vDisk reads, writes and network
- Local vDisk to remote network server location to stress vDisk reads and network
- Remote network server location to local vDisk to stress vDisk writes and network
- Remote network server location to remote network server location to stress the network

In all of these test cases local vDisk reads and writes signify reads and writes to the local system drive (C :\). This drive resides on the server and local reads are handled by the streaming server. The streaming server reads from the shared vDisk image residing on its local RAID 1 volume to satisfy client read requests. Client local writes are performed by writing to the temporary write cache on the server's RAID 10 volume. Clients utilize local CPU, GPU and memory for all operations and only client disk access exercises the server. Each client's local utilization of CPU, GPU and memory is not significant as compared to the impacts of local disk access on the server. Significant factors are the response times to clients and server utilization when a multitude of clients are accessing the disk on server. Hence the client operation tests were primarily focused on simulating the local disk access scenarios and measuring the resulting stress on server.

Results

Standard Mode Boot Characterization

During the boot characterization testing, clients were simultaneously booted from an ACPI S5 (off) power state using Wake on LAN. Booting from this power state eliminates the possibility of client side caching, which can materially impact the test results. The two main entities characterized were client experience and server utilization.

Client Experience

Client experience was characterized through the tracking of two data points. The first is boot time, which is the metric provided by the streaming software that signifies the time elapsed from first network contact with the server until the client agent starts on the client in Windows. The chart in Figure 3 illustrates this information. As expected, boot times were seen to increase as the number of simultaneous clients increased. For comparison purposes, a similarly configured OptiPlex 755 was booted to the local hard drive using the streaming software boot from local hard disk option. The boot time measured using the same metric was 27 seconds.

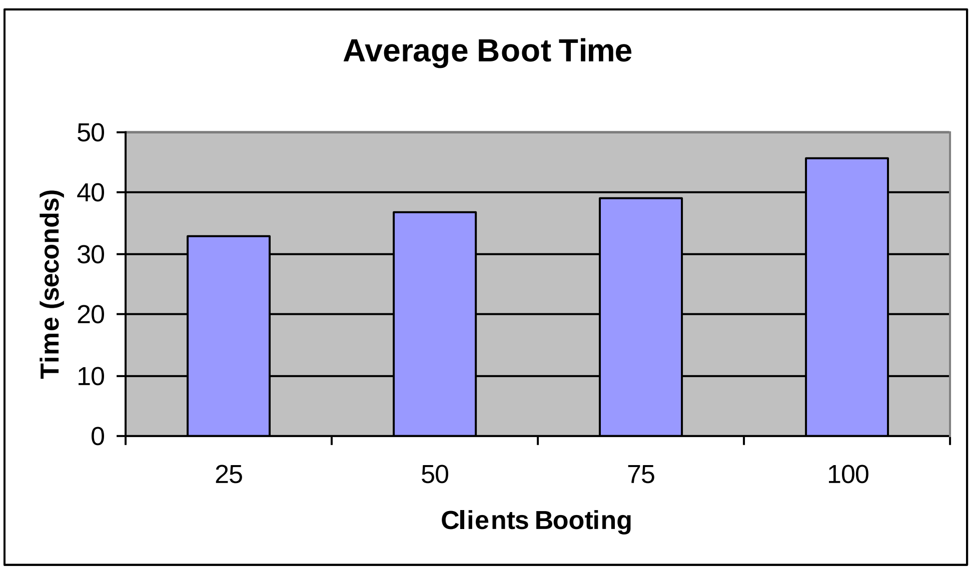

Figure 3: Client boot time

The second metric that comprises the client experience is session retries. This value is provided by the streaming software and signifies the number of times the client did not receive a response to a UDP packet it sent to the streaming server. This occurs every time the UDP packet exceeds its timeout value of 200 milliseconds. When this happens, the client logs a retry and resends the packet to the server. The end user would notice this event in the way of decreased responsiveness of the operating system, or unresponsiveness in extreme cases. The total number of times this occurred in the 6 minute "boot storm" was logged and the result is charted in Figure 4.

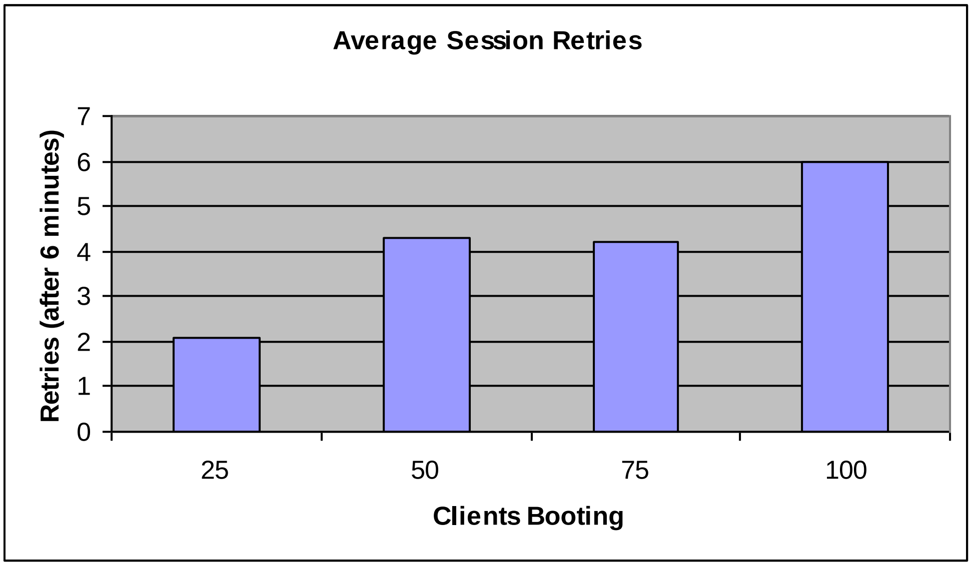

Figure 4: Client session retries

Figures 3 and 4 illustrate that the test environment from this characterization was able to scale to 100 clients booting without a noticeable impact to client end user experience. The increase in boot time seen in figure 3 from 25 to 100 clients of about 15 seconds was a small enough increase that an end user would have no cause to suspect problems. Retries for all numbers of clients tested were sufficiently low that at boot time, there was no noticeable impact to the responsiveness of the clients.

Server Utilization

Server utilization while the clients were booting was measured through tracking the utilization of four major system resources - CPU Utilization, Used Memory, Network Bandwidth and Disk Queue Length. These values provide feedback on the overall server resource utilization during testing.

CPU Utilization

The CPU utilization, tested with a single Intel Xeon 5160 (Dual-core 3.0 GHz) processor, shows a trend of increased CPU load as more clients are booting, with peaks to 100% starting with 75 clients booting (see Figure 5). When clients boot simultaneously the server incurs a processing load 1) to deliver the bootstrap via TFTP services and 2) to deliver the portion of the vDisk image required for successful initial OS boot. As seen in Figure 5, the server CPU utilization increases during the simultaneous client boot and then gradually drops after the delivery is complete. The nature of this CPU utilization increase depends upon the number of clients booted simultaneously. Figure 5 shows that the utilization temporarily peaks to about 85% with 50 clients booting, and this temporary peak increases to 100% when 75 or 100 clients are booting. The duration of CPU usage is also higher with 100 clients booting compared to 50 or fewer clients booting.

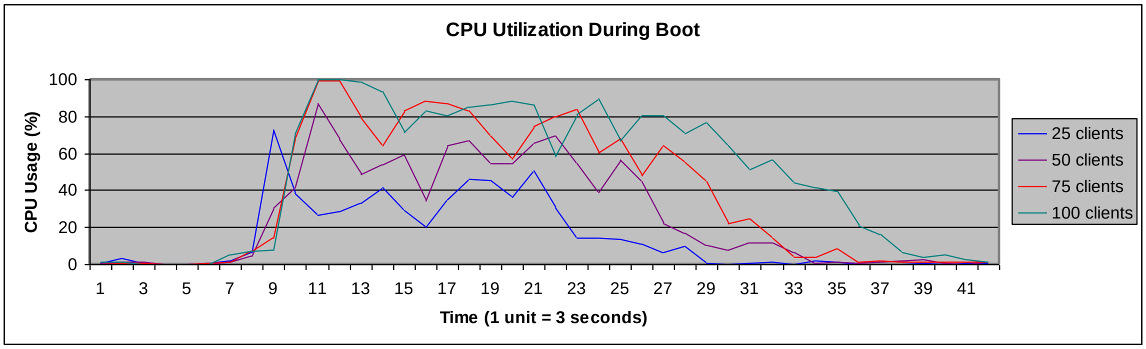

Figure 5: CPU Utilization

The CPU was identified from the test results as one of the more heavily utilized resources on the server during boot testing. The CPU utilization is at acceptable levels for 50 clients at an average of approximately 60% and peaking at 85% utilization. For deployments with a larger number of clients booting simultaneously, a higher speed CPU or a second CPU may be required on the server for nominal utilization levels. It is possible that the resultant server configuration could impact the operations of other measured components.

Memory Utilization

The amount of memory used out of the 4 GB server installed memory is represented in Figure 6. As seen in the CPU graph, memory usage increases steadily as more clients are booted, and it can be seen from the graph that the memory usage decreases after the boot storm has passed.

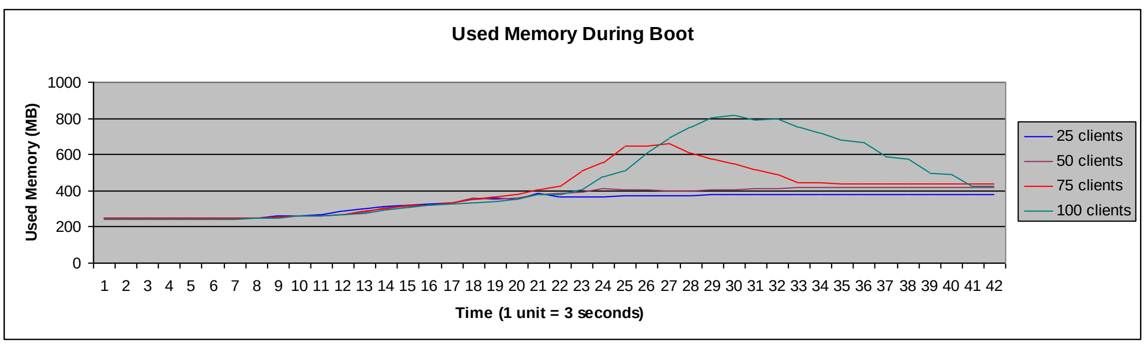

Figure 6: Used Memory

However, even with 100 clients simultaneously booting the server maintained well over 3 GB of memory available to the system. Because of this, memory was not found to be a bottleneck in this solution.

Network Utilization

The outbound network bandwidth was measured to track network utilization on the server. During boot clients perform local disk reads/writes which are translated as network calls to the server. The nature of the inbound network traffic to the server due to these requests is relatively small in bandwidth consumption. The outbound network traffic from the server is significantly larger to satisfy these requests. Figure 7 shows outbound bandwidth on one of the two server onboard gigabit network cards. The test setup utilized both onboard network interfaces on the server, and the bandwidth usage was consistent between the two when the server was under load during boot testing. The usage peaks at about 98% utilization of maximum bandwidth on each interface when booting 75 to 100 clients. As seen in the previous charts, there is an increase in resource utilization as additional clients are booted. This is one of the server resources that are fully utilized in the test setup with 100 clients.

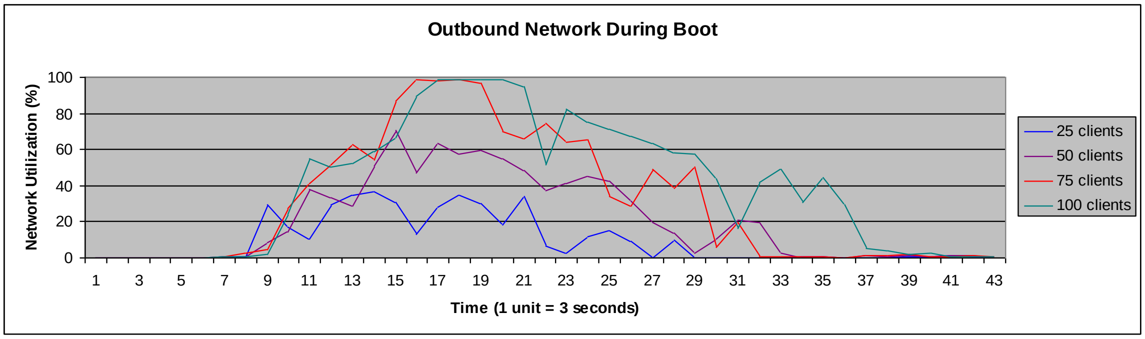

Figure 7: NIC Utilization

For streaming from only one network card, the utilization can be expected to double or even more as compared to the values seen in Figure 7. Since utilization peaks at over 60% when booting 50 or more clients, utilizing only one of the two network interfaces would saturate the connection during a boot storm of 50 clients or more. This may cause other measured factors to increase as well, but was not tested for this solution.

Server Storage

The vDisk image file in the test setup resides in the system drive on the server along with the OS and applications. This is a two disk RAID 1 volume, and client local reads are satisfied from the vDisk on this volume. The temporary write cache for clients is hosted on a separate four disk RAID 10 volume on the server. Client write requests are stored on this write cache volume. Both RAID configurations utilize 300 GB 10K RPM SAS onboard drives.

During client boot testing, it was observed that the client reads generated minimal stress on the RAID 1 system volume, while client writes generated significant stress on the RAID 10 write cache volume. The disk queue length of the write cache volume is represented in Figure 8. Disk queue length is one of the best ways to monitor a server I/O subsystem by showing the average number of write requests that were queued for the disk during the test. This counter includes both queued and in-service requests.

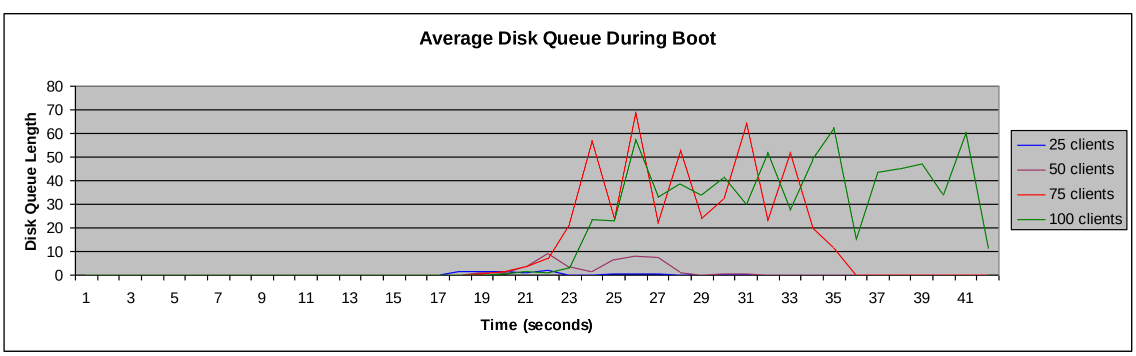

Figure 8: Average disk queue length

From these results, it can be seen that the storage system is the primary server resource gating performance in this solution. If this value exceeds the number of physical disks in the array by twofold it indicates the system is likely developing a bottleneck. With 4 physical disks in this array, 8 is the number that the disk queue would need to be below to indicate no bottleneck at all. When booting 50 or fewer clients, the average disk queue stays at or below 10. But when 75 or more clients are booted, the disk queue length remains between 20 and 70 during the boot storm, and decreases to acceptable levels once the boot storm is complete.

Although the disk queue length appears to be a bottleneck, the actual end user experience impact was negligible. The clients did boot almost 10 seconds slower with an increase in retries from 2 to 6, but the end user experience did not become unresponsive, show an appreciative delay time to detract from end user experience, or cause functionality issues within Windows. For deployments with more than 50 clients booting simultaneously, adding more disk drives to the RAID volume would increase throughput of the server storage and consequently lower disk queue lengths. Another option would be to cache client writes to an alternate location. This could increase load on other server components such as CPU and impact overall performance, depending on modifications made.

Standard Mode Runtime Characterization

As described in the test process section, clients were booted into the OS and then divided into groups which performed different tasks. One group would copy a large file (512 MB) with various source and destination locations while the other group would reboot. The file copies were from local disk to local disk (stressing the vDisk and network), from local disk to remote network share (stressing vDisk reads and network), from remote network share to local disk (stressing vDisk writes and network), and from remote network share to another remote network share (stressing network).

In typical deployments, user data should be stored on appropriate network shares and not on the client local system drive. This is because the local client writes are temporary and will be destroyed at the end of the boot session. All local vDisk writes in the test scenarios below can be interpreted as simulating the temporary write activity to the system drive performed by OS and applications. Examples of temporary write activity or write caching by applications include the temporary space required for Microsoft Internet Explorer operations or for Microsoft Office applications. A single large continuous write operation of this size for temporary OS or application operations can be considered an extreme case relative to normal day-to-day client usage. Additionally, multiple clients performing this local write while other clients boot represents an unusually intense use case. The primary intention of simulating such a local write activity is to understand the impacts to the solution of worst case extreme stress scenarios. The results of the test scenarios in the following sections should be interpreted in this context.

Remote write operations in the tests below simulate the data save operations to a network share for user and application data. This is characterized in the tests by copying a large 512 MB file to a remote network share. Again multiple clients perform this remote write while other clients boot.

Client Experience

The chart in Figure 9 shows boot times during two test cases, 25 clients copying while 25 reboot, and 50 clients copying while 50 reboot. The different bars represent the different copy source/destination locations from the tests as described earlier.

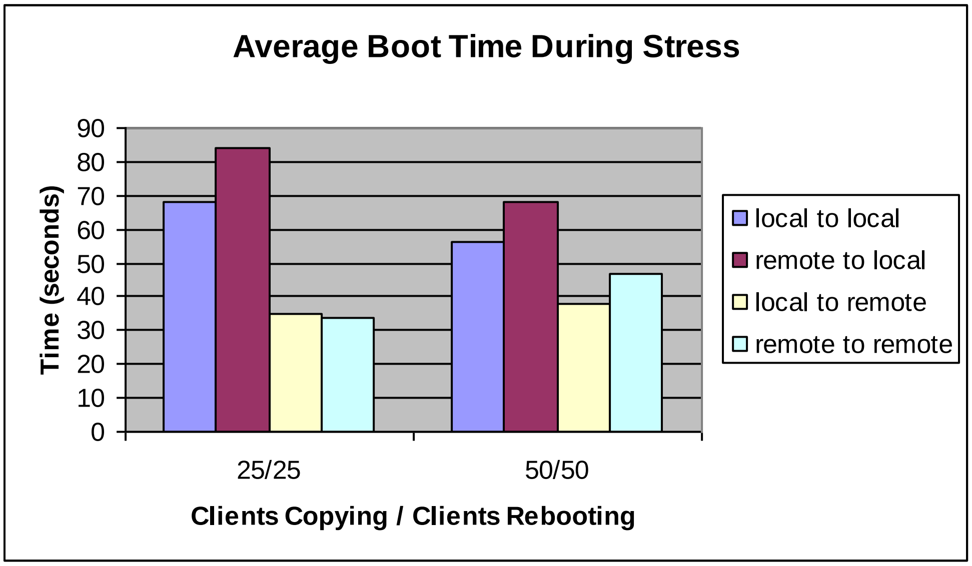

Figure 9: Client boot time

From these results, a differentiation is seen between tests incurring a large number of vDisk write operations and those that did not. Both "local to local" and "remote to local" had significant increases in boot times. These two vDisk write intensive scenarios show a minimal differentiation as compared to the non-write-intensive scenarios.

The chart in Figure 10 shows client retries from the same two test cases. Retries are much higher for vDisk write-intensive operations, and increase significantly when doing the same operations from the 25/25 test to 50/50.

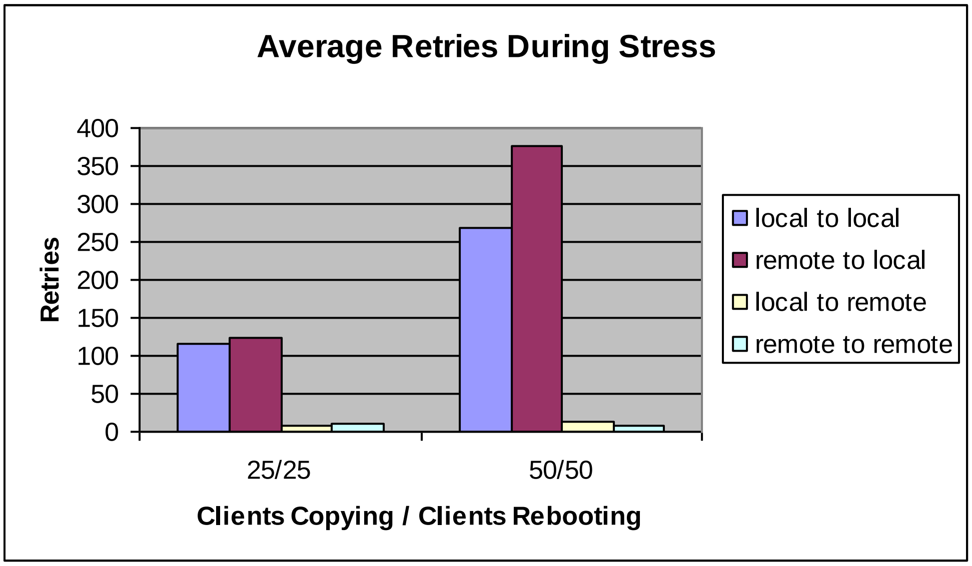

Figure 10: Client session retries

With both test results considered, a trend develops with a decline in performance when doing vDisk write-intensive operations. In the tests with no significant amount of vDisk writes performance degradation from boot testing is almost nonexistent. In situations where large files are being constantly transferred to the local client's vDisk, performance did suffer in the number of retries.

This problem can be alleviated by having network shares set up and mapped for all users and encourage no files to be saved to the local drive (C:\), both for performance reasons and for the fact that no files saved to the local drive are retained after reboot with this solution. Local vDisk writes that occur when the OS and applications write temporary data are usually smaller in nature.

Server Utilization

Server utilization was characterized with the same four data points as in the boot testing; CPU Utilization, Memory Utilization, Network Bandwidth and Disk Queue Length. The values for the 50 client copy / 50 client reboot test have been charted below for these server components. This test is chosen for analysis since it is the most extreme scenario among the four copy/reboot tests.

CPU Utilization

The CPU utilization during the copy/boot stress tests is shown in Figure 11.

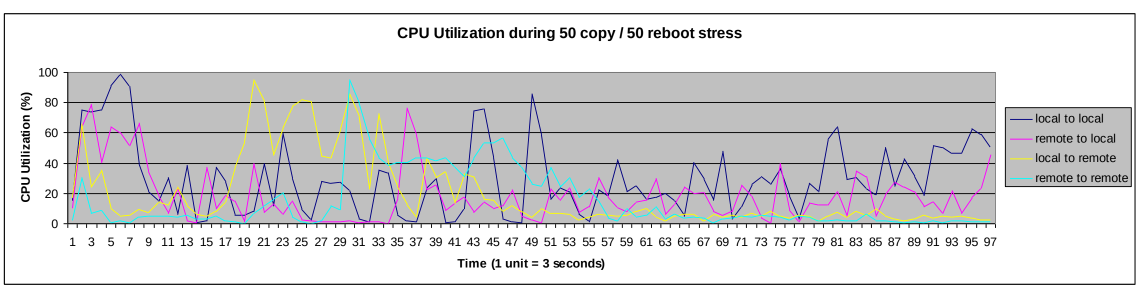

Figure 11: CPU Utilization

The CPU utilization chart shows higher sustained utilization on copy/boot tests versus the previous boot only tests. This can be attributed to the larger number of operations being performed by the server with the combination of file copying and client boots. As seen with the client experience results, CPU has higher utilization levels for tests which are vDisk write-intensive. Even though these tests have higher sustained usage, all tests did have peaks of similar amplitude. This shows that the CPU usage at peak is similar for all test scenarios even though some do not return to lower levels as quickly.

Memory Utilization

Memory usage is also higher in these test cases than in boot only scenarios as seen in Figure 12. A similar trend as above can be seen with memory usage being higher on vDisk write-intensive tests.

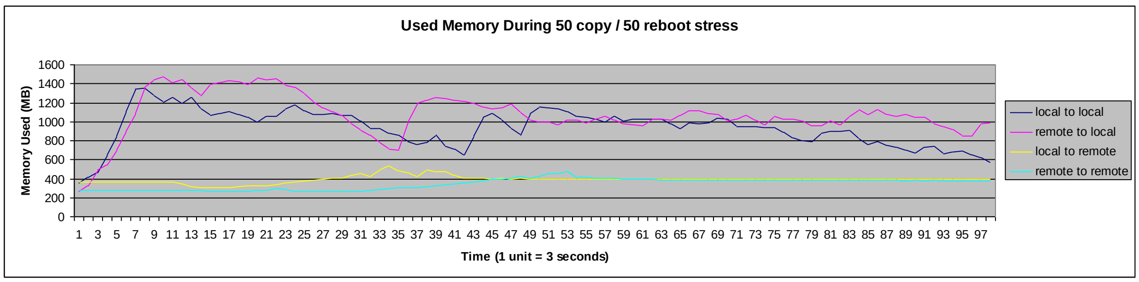

Figure 12: Used Memory

Memory usage remains significantly lower than the total memory installed in the server. Even during the time of highest memory usage in stress testing, memory used never exceeds 40%.

Network Utilization

Outbound network bandwidth for operation tests are shown in Figure 13. As mentioned in the earlier section, the outbound server data traffic is much higher than inbound client request traffic. As with boot only testing, these results are from one of the two onboard server NICs and streaming usage would increase if only one NIC is used.

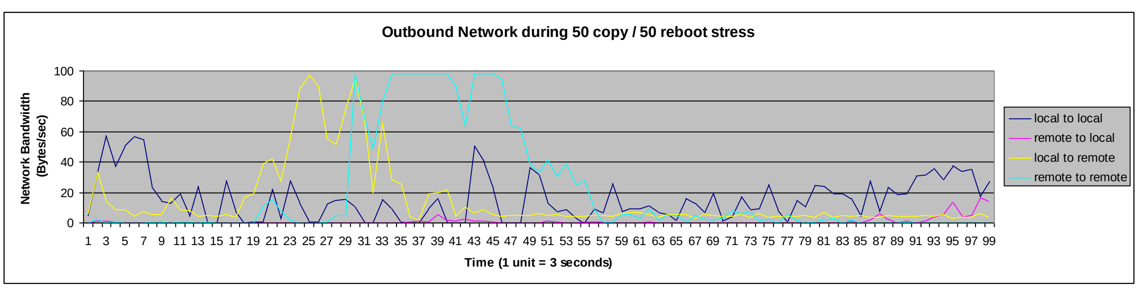

Figure 13: NIC Utilization

The tests which perform remote copy operations, operations that are not write-intensive to the local drive, exhibit much higher network throughput. This can be explained by the fact that copying to a remote location not on the server removes any local server storage bottleneck and allows network transfers to run at their maximum throughput (see Figure 13). Peaks in these remote file copies are about 98%, which is similar to the previous boot only testing results.

Server Storage
During client operation testing, the disk subsystem observation was similar to that of boot only testing. The client reads generated minimal stress on the RAID 1 system volume, but client writes generated significant stress on the RAID 10 write cache volume. The disk queue length of the write cache volume is represented in Figure 14. The chart shows the average disk queue length of I/O activity during client operation tests. As in other operation tests, the write intensive tests show significantly different resource utilization than non write intensive tests.

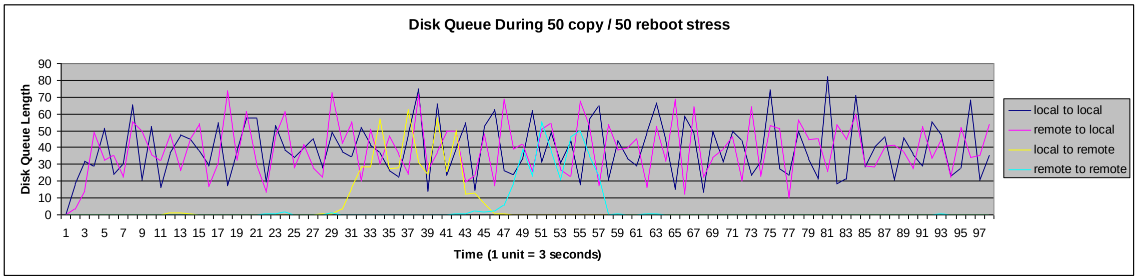

Figure 14: Average Disk Queue Length

As seen in the previous boot only test results, the storage subsystem is the primary server resource limiting performance in this solution. For non-write-intensive tests to the vDisk, the average disk queue briefly peaks to 60 and tails back off, but for write-intensive vDisk tests, the disk queue length remains between 20 and 70. Compared to the boot only testing, the disk queue lengths stay higher for much longer with a corresponding increase in I/O response times. These values are significantly higher than the normal acceptable disk queue length value of twice the number of physical drives in the RAID set, which is 8 in this case. The higher disk queue lengths are associated with the extreme nature of the stress exerted by the write operations. Such a high local write operation would not be incurred usually for temporary client OS and application caching. Nevertheless this characterization provides disk utilization data for worst case stress scenarios of client local write operations. The client write requirements of deployments should be assessed carefully and the server disk subsystem should be planned appropriately.

Even though the disk queue length stayed high for a significantly longer time, the clients remained responsive and did not have any issues during the testing. Adding more physical drives or faster drives should help bound the disk queue length values but may have other impacts.

High Availability Observations

Similar characterization was performed on a High Availability (HA) configuration using the servers listed in the Test Infrastructure section above. The same 100 clients were used for this characterization, and the same tests were performed. The HA configuration moves all vDisks, write caching, and database activity from a 4 disk RAID 10 on the streaming server to an 8 disk RAID 10 on the NAS storage server. The storage server handles all vDisk reads, client writes, and database activity, leaving the streaming servers to only manage the OS streaming itself. Another benefit of moving the data off of the streaming servers is it allows clients to fail over to another streaming server if the server they are currently using fails.

Client Experience

Figure 15 shows the client retries when booting 100 clients with both a standalone server, and an HA configuration.

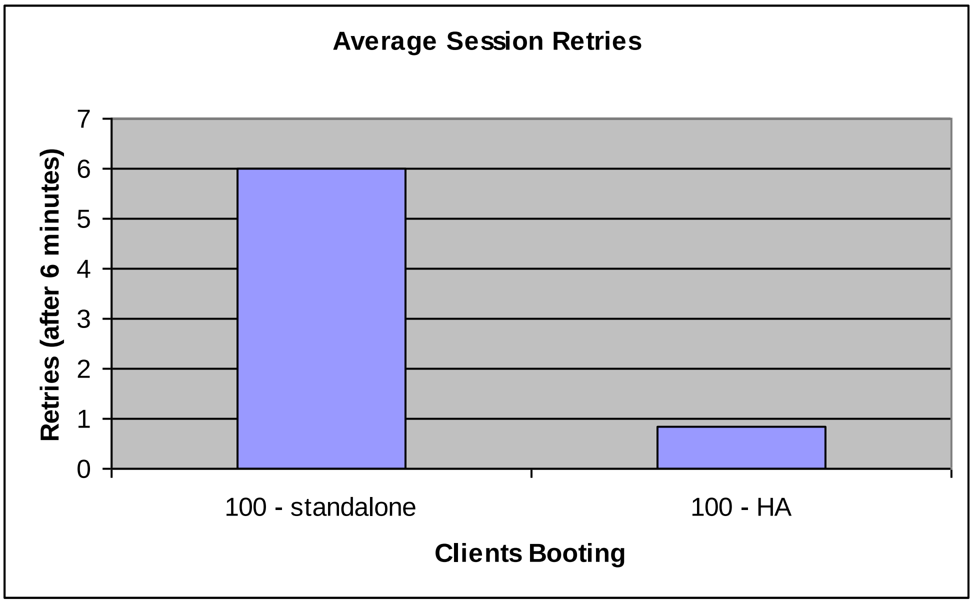

Figure 155: Average Session Retries

As you can see from the chart, the retries significantly dropped when moving to the HA solution with more drives in the RAID 10. Despite this improvement, the user experience was not noticeably affected due to the uncompromised experience in the standalone setup.

Server Utilization

The most significant change from a single standalone server to the HA configuration was server storage performance. Figure 16 illustrates the stress from the standalone server discussed above, compared to the NAS storage server used in HA characterization.

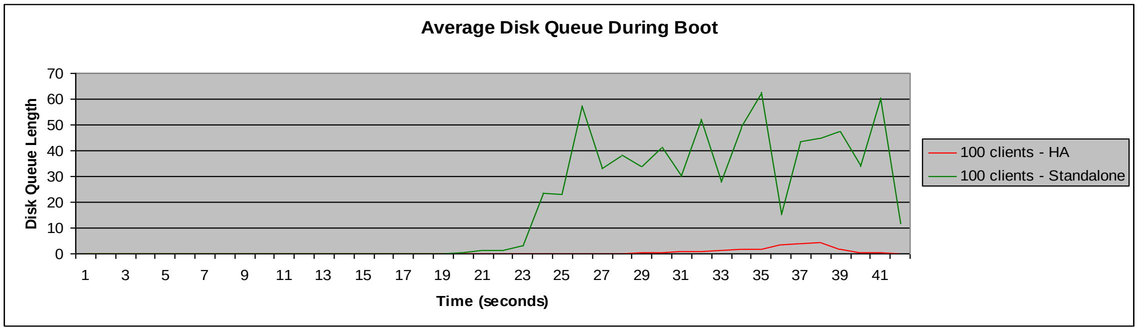

Figure 166: Average Disk Queue Length

By adding more spindles to the RAID, the storage subsystem is able to keep up with the flood of data that occurs during a boot storm without having high disk queue lengths. The chart shows that the stress generated on the RAID 10 volume of the NAS storage server was much lower than that generated on the standalone streaming server.

Conclusion

The performance characterization data discussed in this paper are applicable to the specific configuration listed. The actual performance realized will vary depending on the particular deployment scenario and the environment. Nevertheless, there are some important observations in the test results that are applicable to most deployments.

- Although the standalone server resource utilization levels of the 50:1 client to server ratio conform to traditional enterprise best practices, the 100:1 ratio characterized did not cause observable poor client performance or unresponsiveness. With an HA configuration, 100 clients conform to traditional enterprise best practices.
- Onboard Storage: The onboard storage in the PowerEdge 2950 was adequate for all test cases covered in this characterization. The results showed that it was the most utilized component of the server and a candidate for upgrade if conformance to traditional utilization levels is required.
- Networking: All testing described in this paper was performed with gigabit networking. In an end-to-end 100 megabit environment, performance may be expected to degrade perceptibly depending on the usage trends of the users in that environment. As tested, the end-to-end gigabit networking was sufficient for all characterizations, but 75:1 (or higher) client to server ratios did cause the server to temporarily peak at levels higher than traditional network utilization best practices.
- Processor: The single Intel Xeon 5160 used in this testing was adequate for acceptable client experience in all tests performed. Acceptable server CPU utilization levels should be defined for deployments and the CPU speed, number of cores, or CPU count should be scaled appropriately to maintain traditional utilization levels.
- Server Memory: Server memory utilization was reasonably low in all test case scenarios. The available 4 GB system memory was sufficient to handle 100 client boots and operations.
- vDisk Access: During client operation testing, there was a significant difference in performance tests which were write-intensive to the vDisk and those which were not. If the users of this solution are not performing tasks which are vDisk write-intensive, performance can be expected to degrade less in high load situations.

Appendix A

System configuration details for the test components:

Network Switch:

- Spanning tree mode enabled on network switches to resolve loops for multi-port uplink switch configurations
- Port fast mode enabled on network switches

Server Storage:

- Write-back cached mode enabled for the RAID 10 container in the firmware of the Dell Power Edge RAID Controller (PERC 5/i). This enables caching data writes to the disk drives and improves write I/O performance. Data not yet committed to the disk is backed up via a battery on the PERC.

Server Operating System:

- OS registry key KEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Session Manager\MemoryManagement\LargeSystemCache set to 1 to assure the maximum amount of 960 MB RAM is available for file system disk caching. Note that on Microsoft 2003 Server R2 Standard Edition (32-bit), the maximum supported cache size is 960 MB.

THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.

Dell, PowerEdge, PowerVault, PowerConnect, OptiPlex, and the Dell logo are trademarks of Dell Inc. Other trademarks and trade names are the property of their respective owners and Dell disclaims proprietary interest in the marks and names of others.

© Dell Inc. 2007. All rights reserved. Reproduction in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell. Information in this document is subject to change without notice.