HP StorageWorks Enterprise Virtual Array online firmware upgrade best practices

Part number: 5697–6388
Second edition: January 2007

invent

Contents

Figures

Tables

About this guide

This guide includes the following information:

- Understanding how the EVA online firmware upgrade process works.
- Assessing your SAN environment to determine if it can support an online upgrade.
- Determining how long the storage system will be unavailable during the online upgrade process. This time is required to perform the controller resynchronization.
- Scheduling the best time to perform the upgrade.

NOTE:

This document does not include the detailed procedural steps performed during an online firmware upgrade. For detailed procedures, see *HP StorageWorks 3000/5000 Enterprise Virtual Array updating product software guide* or *HP StorageWorks 4000/6000/8000 Enterprise Virtual Array updating product software guide*, depending on the array model. See Related documentation.

Intended audience

This guide is intended for personnel involved in planning or performing a firmware upgrade on an EVA3000/5000 or an EVA4000/6000/8000. To perform an online upgrade, you should have experience in the following areas:

- Evaluating the storage system configuration to determine its impact on how long the storage system will be unavailable during the online upgrade process.
- Using HP Command View EVAPerf to monitor and record storage system I/O load.
- Identifying operating system and application time out values for all hosts accessing the storage system. And changing the time out values if necessary to support the online upgrade.

Related documentation

The following documents provide related information:

- *HP StorageWorks Command View EVA user guide*
- *HP StorageWorks Command View EVA release notes*
- *HP StorageWorks 3000/5000 Enterprise Virtual Array updating product software guide*
- *HP StorageWorks 4000/6000/8000 Enterprise Virtual Array updating product software guide*

You can find these documents from the Manuals page of the HP Business Support Center website:

http://www.hp.com/support/manuals

For HP Command View EVA, click **Storage software** under Storage, and then select **HP StorageWorks Command View EVA Software** under Storage Device Management Software.

For EVA, click **Disk Storage Systems** under Storage, and then select **HP StorageWorks 3000 Enterprise Virtual Array**, **HP StorageWorks 5000 Enterprise Virtual Array**, or **HP StorageWorks 4000/6000/8000 Enterprise Virtual Arrays** under EVA Disk Arrays.

Document conventions and symbols

Table 1 Document conventions

Convention	Element
Blue text: Table 1	Cross-reference links and e-mail addresses
Blue, underlined text: http://www.hp.com	Web site addresses
Bold text	• Keys that are pressed • Text typed into a GUI element, such as a box • GUI elements that are clicked or selected, such as menu and list items, buttons, tabs, and check boxes
Italic text	Text emphasis
`Monospace` text	• File and directory names • System output • Code • Commands, their arguments, and argument values
`Monospace, italic` text	• Code variables • Command variables
`Monospace, bold` text	Emphasized monospace text

△ CAUTION:

Indicates that failure to follow directions could result in damage to equipment or data.

IMPORTANT:

Provides clarifying information or specific instructions.

NOTE:

Provides additional information.

TIP:

Provides helpful hints and shortcuts.

HP technical support

Telephone numbers for worldwide technical support are listed on the HP support web site: http://www.hp.com/support/.

Collect the following information before calling:

- Technical support registration number (if applicable)
- Product serial numbers
- Product model names and numbers
- Error messages
- Operating system type and revision level
- Detailed questions

For continuous quality improvement, calls may be recorded or monitored.

Documentation feedback

HP welcomes your feedback.

To make comments and suggestions about product documentation, please send a message to storagedocs.feedback@hp.com. All submissions become the property of HP.

1 Understanding the online firmware upgrade process

During the life of an EVA storage system, it will be necessary to periodically upgrade the system firmware. Upgrading to the most current firmware release ensures that the storage system benefits from ongoing design improvements and enhancements.

Performing the firmware upgrade online minimizes the impact on the hosts and applications accessing the storage system. It is not necessary to halt I/O or suspend operation of the hosts. Because of the significant benefits offered by the online upgrade, every effort should be made to use this feature when upgrading the storage system firmware.

Online upgrade can be performed on all operating systems supported by the EVA. In addition, online upgrade has been tested successfully with applications such as Oracle, Microsoft SQL Server, and Microsoft Exchange.

Definition of terms

The following terms are used throughout this document and are defined as follows:

- **Controller resynchronization**: The interval during which the storage system controllers are unavailable to process host I/Os. The SAN environment must be able to accommodate this interval.
- **Offline upgrade**: An upgrade process that requires stopping all host I/O to the storage system. This typically requires all applications and hosts accessing the storage system be shut down prior to the firmware upgrade. The hosts and applications are restarted after the upgrade is complete. The EVA supports this type of upgrade.
- **Online upgrade**: An upgrade process that enables the storage system controller firmware to be upgraded while still receiving host I/Os. The firmware on both EVA controllers is upgraded simultaneously, and the resultant resynchronization makes the storage system inaccessible for a brief period of time. The resynchronization represents a single disruption to I/Os.

How online upgrade works

The following provides an overview of the major steps involved in the online upgrade process.

1. Download the new firmware image to the controllers.
2. Validate the firmware image and copy it to flash memory.
3. In an HP Continuous Access EVA environment, the peer storage system is notified that a resynchronization will occur.
4. Resynchronize the controllers.
5. Host I/Os not processed by the controllers during the resynchronization is retried by the host and completed.

Processing of host I/Os continues while the firmware images are downloaded and validated, and while HP Continuous Access EVA systems are being notified. In most situations, the controller resynchronization interval is less than 40 seconds, which is well below the 60 second timeout allowed by most operating systems.

Online upgrades in HP Continuous Access EVA environments

The following recommendations apply when performing an online upgrade in HP Continuous Access EVA environments:

- The HP Command View EVA management server is often connected on the fabric between the source and destination storage systems in such a way that one storage system is considered to be local to the management server. That is, management commands do not pass across the intersite link (ISL). An online upgrade should be performed from a local management server to avoid sending the firmware code over the ISL. The transfer of the firmware image across the ISL can cause dropped jobs and degraded replication performance.
- It is possible to perform online upgrades while replicating data. Table 2 indicates the configurations in which replication must be suspended during the upgrade. For example, if a storage system is both a source and a destination for synchronous replication, the replication groups for which the storage system is the destination should be suspended. The groups for which the storage system is the source do not need to be suspended.

Table 2 Suspending replication during an online upgrade

Environment	Synchronous replication	Asynchronous replication
The storage system being upgraded is an HP Continuous Access EVA source	Replication does not need to be suspended	Replication does not need to be suspended
The storage system being upgraded is an HP Continuous Access EVA destination	Replication should be suspended to reduce the probability that the resynchronization on the destination impacts workload on the host.	Replication does not need to be suspended

Online disk drive firmware upgrades

New disk drive firmware is typically included with a new XCS or VCS release. To ensure optimal storage system performance, HP recommends that you upgrade the disk drive firmware when upgrading the controller firmware. If necessary, you can delay upgrading the disk drive firmware, but the controller firmware should be upgraded online immediately as soon as possible. Controller firmware upgrades provide the primary benefits and enhancements of a new firmware release. The disk drive firmware updates can be done at a later date.

With the release of HP Command View EVA 6.0.2, online firmware upgrade of multiple disk drives is supported. This enables you to perform an online upgrade of all storage system components in one operation. In general, performing the upgrade during a period of low I/O activity will help ensure the success of the upgrade. It is recommended to adhere to the same parameters as outlined for the XCS firmware process while performing the online disk firmware upgrade.

NOTE:

If you choose to delay upgrading the disk drive firmware, there will be controller events indicating that out of date disk drive firmware is being used. These events will be disappear when the disk drive firmware is updated.

2 Evaluating the SAN for an online upgrade

Before performing an online upgrade, the SAN environment must be evaluated to determine system timeouts, identify I/O patterns, and select the best time to perform the upgrade. This will help ensure the success of the online upgrade.

Managing host I/O timeouts

The default values for host operating parameters such as LUN timeout and queue depth are typically set to values that ensure proper operation with the storage system. These values are appropriate for most storage system operations, including online firmware upgrades. In general, host LUN timeouts of 60 seconds or more are adequate to accommodate an online upgrade. In most situations it will not be necessary to alter these settings to perform an online firmware upgrade.

If any host timeout values have been decreased to less than 60 seconds, it will be necessary to reset them to their original default value. The following section provide a summary of the steps and commands involved in checking and changing timeout values for each supported operating system. Refer to the operating system documentation for more information.

NOTE:

- All operating systems are supported for online controller firmware upgrades when the following guidelines are followed for both raw and file system devices.
- Unless otherwise indicated, cluster testing has not yet been completed in combination with online controller firmware upgrades. Cluster information will be added to this document when the testing is complete.

Microsoft Windows

NOTE:

- Ensure that MPIO is installed and both paths are active and have access to the array.
- Boot devices are supported with online controller firmware upgrades.

Checking timeout values

- The key timeout setting is in the registry at `HKEY_LOCAL_MACHINE/SYSTEM/CurrentControlSet/Services/Disk/TimeoutValue` is set to 3c (60 seconds)

HP-UX

CAUTION:

If Serviceguard is being used, the timeout value is 45 seconds.

Default timeout values

- Sdisk timeout: 30 seconds
- (LVM) lvol timeout: 0 seconds

Checking timeout values

- Device within a LVM volume group: `pvdisplay /dev/dsk/cxtxdx`
- lvol physical volume with an LVM group: `lvdisplay /dev/vgxx/lvolx`

Changing timeout values

```
Pvchange -t (seconds) /dev/dsk/cxtxdx
```

Linux

The following configuration recommendations apply to both RedHat and SUSE.

QLogic Driver parameters (failover enabled):

- qdepth = 16
- port_down_retry_count = 30
- login_retry_count = 30
- failover = 1
- load_balancing = 1
- excludemodel = 0x0
- auto_restore = 0xA6

QLogic Driver parameters (single path):

- qdepth = 16
- port_down_retry_count = 64
- login_retry_count = 30
- failover = 1
- load_balancing = 1
- excludemodel = 0x0
- auto_restore = 0xA6

Emulex Driver parameters (single path)

- HPELXLPFC=y
- nodev_timeout=60
- qdepth=30
- discovery_threads=1

To check or set Linux parameters, use the `set_parm` executable found in the `/opt/hp/<driver-specific>` directory. When executed, the options to change timeout values are displayed.

NOTE:

- The timeout values must be increased for Emulex single path (without multipath support) and QLogic single path being used in the environment. This is not only important for online upgrades but for general data integrity.
- Online controller firmware upgrades are supported with Linux boot devices.

Solaris

Solaris supports online controller upgrades with the following driver timeouts.

- Sun Drivers(qlc or emlxs): 60 seconds
- QLogic (qla2300): 60 seconds
- Emulex (lpfc): 60 seconds

Checking or changing timeouts

For Sun drivers, add the following lines to `/etc/system` file:

```
set sd:sd_io_time = 60
set ssd:ssd_io_time = 60
```

For QLogic, edit the `/kernel/drv/qla2300.conf` file and change the `hbax-link-down-timeout` value to 60 as follows:

```
hba0-link-down-timeout=60;
```

For Emulex, edit the `/kernel/drv/lpfc.conf` file and change the `linkdown-tmo` value to 60 as follows:

```
linkdown-tmo=60;
```

No cluster testing has been performed on Solaris.

Tru64

Tru64 has a 60 second timer for media command timeouts. In most cases, up to 5 retries can occur within the kernel's I/O stack, which results in a 5–6 minute window before the I/O is returned to the file system. No changes should be made to any Tru64 UNIX settings when performing an online firmware upgrade.

Online controller firmware upgrades are supported in Trucluster four-node environments where the EVA is the cluster boot device.

OpenVMS

Online controller firmware upgrades on OpenVMS included testing in 16–node OpenVMS cluster environments where the EVA was the cluster boot device.

In clustered environments where the system disks or quorum disks are on the storage system, the cluster will stop processing requests while the storage system is resynchronizing. The cluster will become available after the resynchronization. For the duration of an online code load, the cluster may log messages related to the inability to access the quorum disk during the upgrade.

In environments where both shadow disks are on the same EVA or where they are on different arrays in a 2–node cluster, activity will stop during the online controller firmware upgrade.

IBM AIX

Checking or changing timeouts

AIX requires the disk settings shown in Table 3 for the native multipath drives.

Table 3 IBM AIX timeout settings

Setting	Value	Description
PR_key_value	NA	Sets the key value for persistent reservations. Persistent reservations are not supported.
Algorithm	fail_over	Sets the load balancing algorithm to fail_over. All I/O uses a single path. The remaining paths are in standby mode. The value round_robin is not supported.
hcheck_interval	60	Sets the path health feature to check each device every 60 seconds.
hcheck_mode	nonactive	Specifies the I/O paths monitored by the path health checking feature: nonactive—Checks all I/O paths for Failed status, and checks standby paths for Used/Opened devices.
queue_depth	8	Sets the queue depth.
reserve_policy	single path	Sets the reserve policy to standard SCSI-2 reservations.
rw_timeout	60	Sets the read/write timeout to 60 seconds.

Managing application I/O timeouts

Applications are typically insulated from the online controller firmware upgrade by the operating system and HBA driver software. Consequently, if the application is running on a properly configured operating system, the online controller upgrades will be performed successfully.

It may be necessary to determine if any applications have timeout requirements that are more stringent than those of the operating system. Any applications with more stringent timeouts should be evaluated to determine if the application timeous will be exceeded by the resync interval that occurs during the online upgrade.

Extensive testing has been done to evaluate the impact of an online upgrade on the following applications:

- Oracle 10g
- Microsoft Exchange 2003
- Microsoft SQL Server 2005

Under both light and heavy I/O loads, these applications reported no errors of any kind during the online firmware upgrade. The applications did not fail during the upgrade and returned to normal operation immediately following completion of the upgrade. However, there was an impact on performance. To minimize the performance impact on users, the online upgrade should be performed during a period of low user activity

For detailed testing results, see "Effect of online firmware upgrade on application resiliency" on page 27.

Assessing the impact of I/O load on the upgrade

Because the online firmware upgrade is performed while host I/Os are being serviced, the I/O load can impact the upgrade process. In general, performing the upgrade during a period of low I/O activity will help ensure the success of the upgrade.

There are three primary ways in which I/O load impacts the online upgrade process.

- First is the ability of the storage system to perform the upgrade in parallel with host I/O. These two processes compete for internal controller resources, with host I/O having the higher priority. Therefore decreasing the host load, by finding a period of time where host demand is the smallest, allows a greater number of resources to be used to execute the online controller firmware upgrade as quickly as possible.
- Second is the ability of the storage system to communicate with HP Command View EVA. The first steps of the firmware upgrade use a process that has a lower priority than host I/O. Under heavy I/O load, the time required to perform the upgrade can exceed the timeouts expected by HP Command View EVA for the completion of the upgrade process. Although the firmware upgrade is typically successful, the timeouts may cause HP Command View EVA to lose connectivity to the storage system and can cause confusion in the upgrade process.
- Third is the queuing effect on host I/O. During the controller resynchronization interval, I/Os can continue to queue on the host. This creates a backlog of requests that must be processed by the storage system when the firmware upgrade is complete. Although the resynchronization interval may be less than 60 seconds, a large backlog may require a great deal of time to clear, and some hosts may experience I/O time outs.

As a result of these factors, it is recommended that an evaluation of the workload be done before performing the upgrade. This will help identify periods of low I/O activity during which the upgrade can be performed successfully.

Selecting an appropriate time to perform the online upgrade

To ensure the success of an online upgrade, selecting the appropriate time is essential. Selecting a period of low I/O activity will ensure the upgrade completes as quickly as possible, and will avoid disruptions to host and applications due to timeouts. Attempting to upgrade a storage system that is in the middle of processing a large I/O intensive batch job will likely causes hosts to lose connectivity with the storage system. Consequently, finding an appropriate time for the upgrade is very important.

The number of disk drives in the storage system influences the I/O demands that are tolerable during an online firmware upgrade. HP recommends that during an online upgrade the I/O demand on the storage system be limited to the values shown in Figure 1, Figure 2, Figure 3, and Figure 4. The load values for both the number of IOPS and throughput (MB/s) must be not be exceeded.

For example, an EVA8000 with 168 disk drives should have a demand less than both 70 MB/s and 7,000 IOPS and an EVA5000 with 168 disk drives should have a demand less than both 40 MB/s and 4,000 IOPS. This will ensure that a healthy storage system can resynchronize well with in the defined resynchronization window.

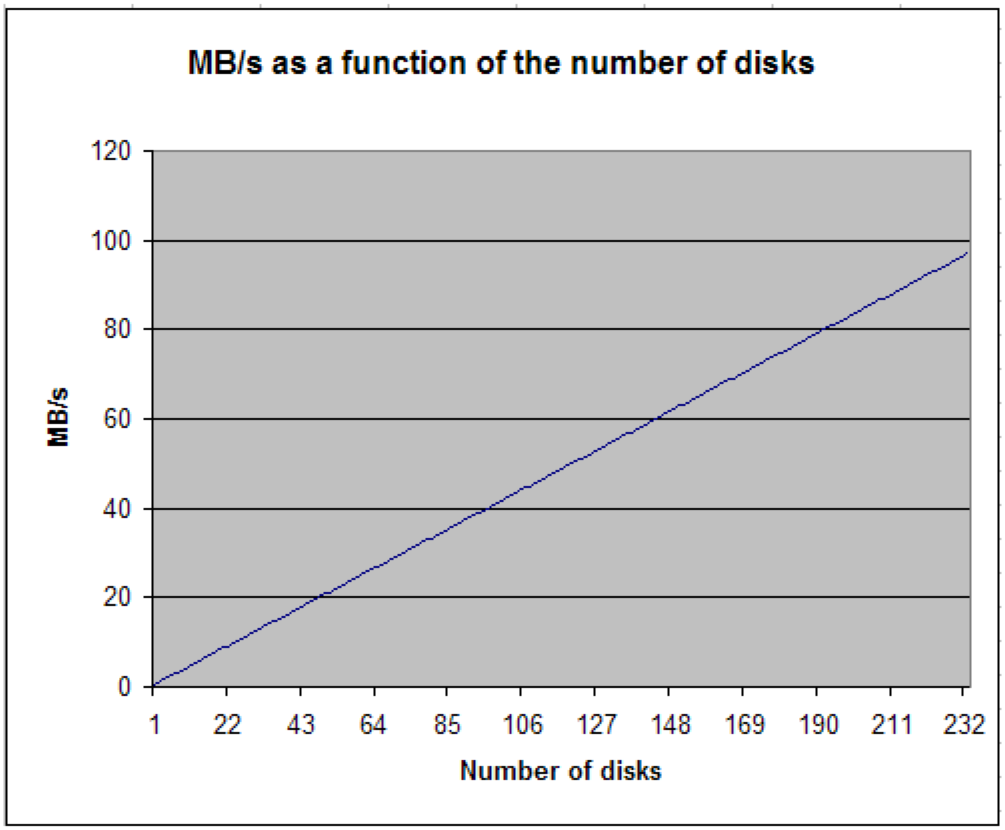

Figure 1 Determining the maximum I/O throughput (MB/s) load for EVA4000/6000/8000

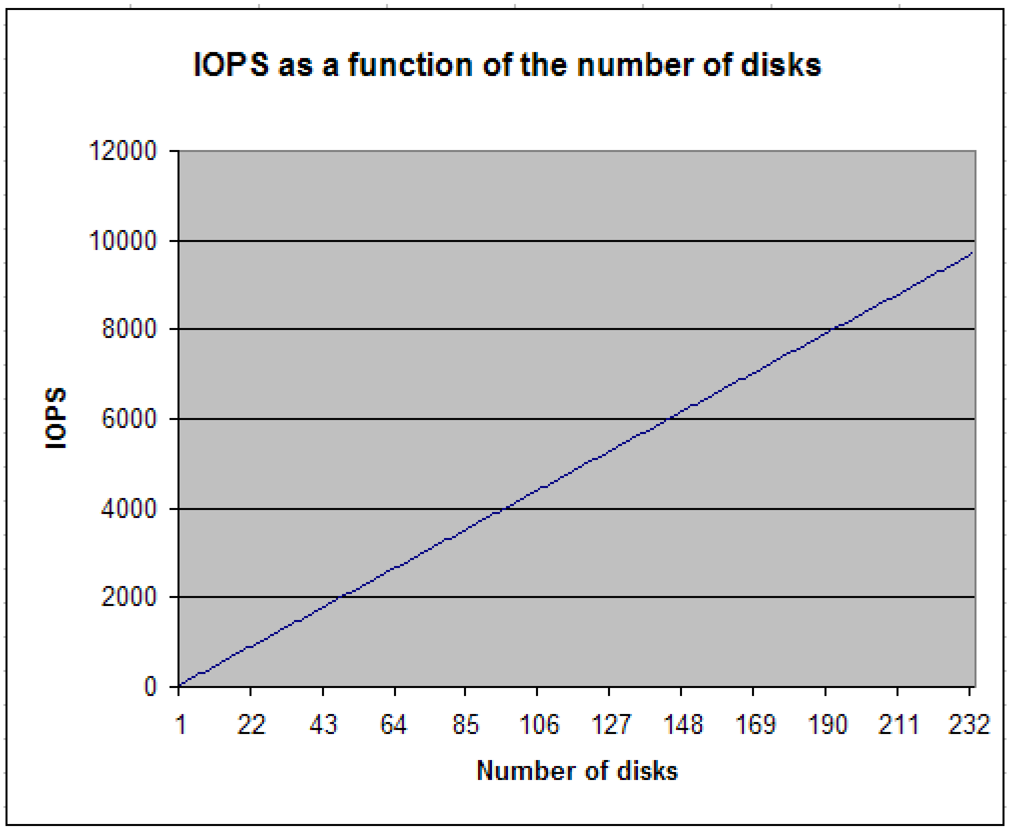

Figure 2 Determining the maximum I/Os per second (IOPS) load for EVA4000/6000/8000

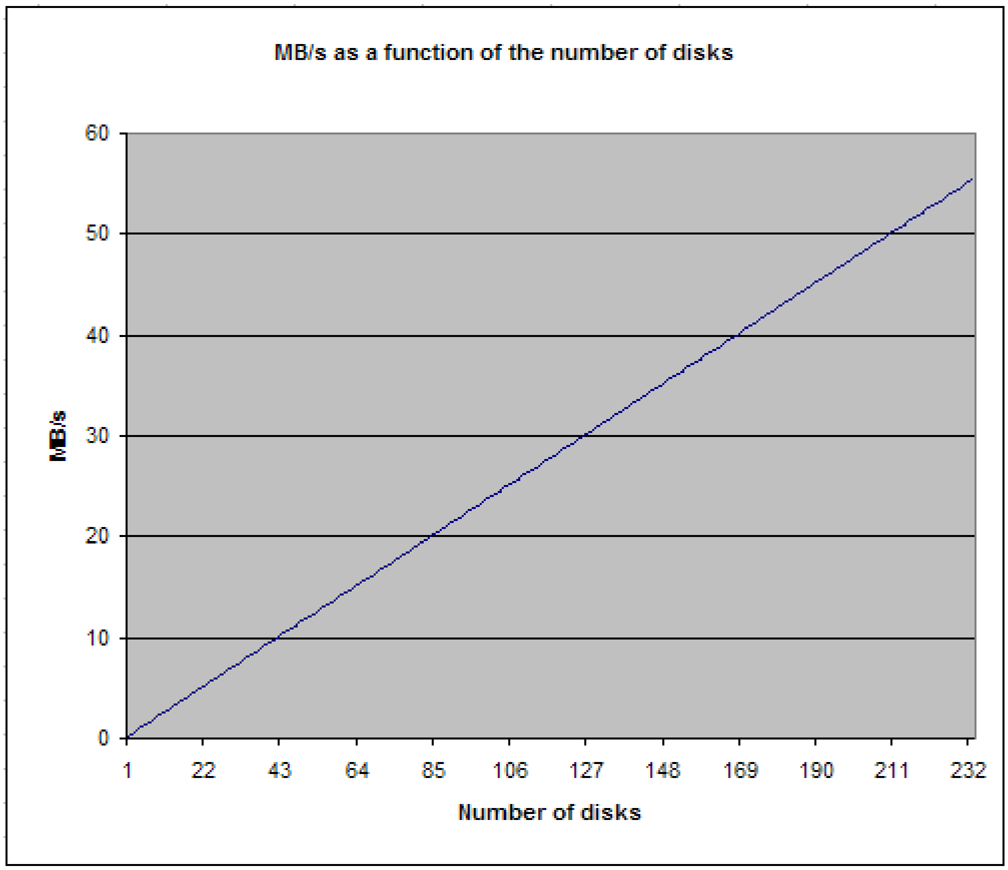

Figure 3 Determining the maximum I/O throughput (MB/s) load for EVA3000/5000

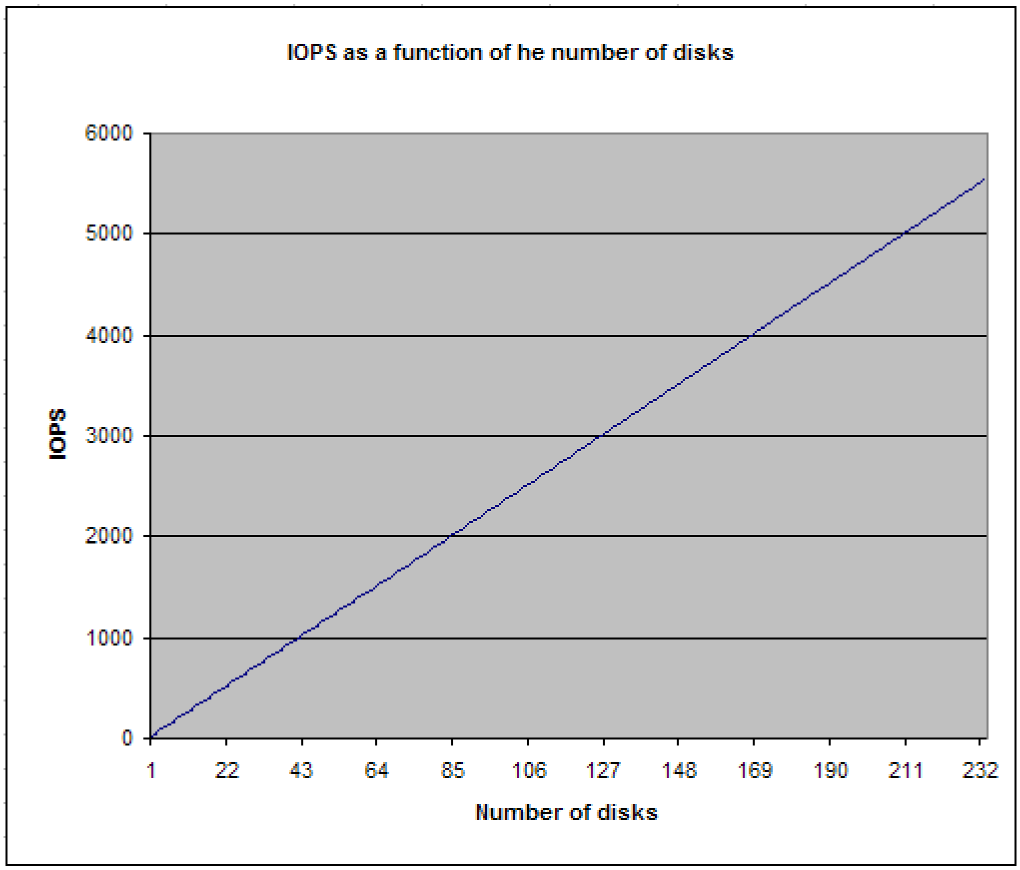

Figure 4 Determining the maximum I/Os per second (IOPS) load for EVA3000/5000

Analyzing storage system utilization using HP Command View EVAPerf

The HP Command View EVAPerf tool can be used to gather and analyze statistics on storage system utilization. This section provides recommendations on using HP Command View EVAPerf to gather statistics to help identify periods of low storage system activity.

If you are not familiar with HP Command View EVAPerf, see the *HP StorageWorks Command View EVA user guide* for more information. You may also want to read through the HP Command View EVAPerf white paper available on the following website:

http://h18006.www1.hp.com/storage/arraywhitepapers.html

The most important statistics for evaluating the state of the storage system for online firmware upgrades are the host port statistics (hps) and the virtual disk statistics (vd). The HP Command View EVAPerf data can be used to calculate the total demand on the storage system by summing the columns for a given point in time.

Capturing host port statistics

Use the following command to capture the host port statistics:

`evaperf hps -csv cont` *`n`* `-dur` *`n`* `-sz` *`array`* `-fo` *`filename`*

- `-csv` specifies comma separated output. This format is useful for importing large amounts of data into a spreadsheet for analysis.
- `-cont` *`n`* runs command continuously at interval specified by n. The default interval is one second.
- `-dur n` specifies the duration of a continuous mode session. For example, if you enter evaperf hps -cont 2 -dur 30, host port data is displayed at two second intervals for a total of 30 seconds.

- `-sz` *`array`* limits data collection to the specified array(s). You must enter at least one array and can use either the storage system WWN or friendly name.
- `-fo` *`filename`* directs output to a specified filename. Include the path information as necessary.

NOTE:

The HP Command View EVAPerf `as` command provides an alternative to the `hps` command for gathering and displaying IOPS and I/O throughput.

Name	Read Req/s	Read MB/s	Read Latency (ms)	Write Req/s	Write MB/s	Write Latency (ms)	Av. Queue Depth	Ctlr	Node
FP1	3877	127.04	0.6	0	0.00	0.0	2	T00K	Cab2
FP2	2786	11.41	6.2	1141	4.67	16.4	146	T00K	Cab2
FP1	489	8.02	16.5	679	11.13	58.1	47	T02K	Cab2
FP2	1186	38.87	15.8	1263	41.39	58.1	91	T02K	Cab2

Figure 5 Sample host port statistics

Capturing virtual disk statistics

Use the following command to capture the virtual disk statistics:

`evaperf vd -csv cont` *`n`* `-dur` *`n`* `-sz` *`array`* `-fo` *`filename`*

The command parameters are described in the preceding section.

Virtual disk statistics are useful for identifying a virtual disk that has significant activity at a given point in time. For example, assume a storage system with multiple virtual disks that are all nearly idle with the exception of a single virtual disk. To perform an online upgrade, it may be desirable to halt the application that is imposing the load on the busy logical drive. This may be preferable to doing an offline upgrade.

ID	Read Hit Req/s	Read Hit MB/s	Read Hit Latency (ms)	Read Miss Req/s	Read Miss MB/s	Read Miss Latency (ms)	Write Req/s	Write MB/s	Write Latency (ms)	Flush MB/s	Mirror MB/s	Prefetch MB/s	
0	1735	113.75	0.4	10	0.66	16.9	0	0.00	0.0	0.00	0.00	110.34	Defau
0	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0	0.00	0.00	0.00	Defau
1	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0	0.00	0.00	0.00	Defau
1	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0	0.00	0.00	0.00	Defau
6	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0	0.00	0.00	0.00	Defau
6	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0	0.00	0.00	0.00	Defau
7	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0	0.00	0.00	0.00	Defau
7	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0	0.00	0.00	0.00	Defau
2	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0	0.00	0.00	0.00	
2	0	0.00	0.0	1834	15.03	8.4	1193	9.77	0.3	9.80	9.77	0.00	
3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0	0.00	0.00	0.00	
3	0	0.00	0.0	0	0.00	0.0	0	0.00	0.0	0.00	0.00	0.00	

Figure 6 Sample virtual disk statistics

Capturing all HP Command View EVAPerf statistics

If you would like to capture all the HP Command View EVAPerf statistics, use the following command:

`evaperf all -csv cont` *`n`* `-dur` *`n`* `-sz` *`array`* `-fo` *`filename`*

The command parameters are described in the preceding section.

Storage system utilization data analysis

When evaluating the performance of the EVA, it is important to distinguish between the two fundamental types of I/O patterns: small block reads/writes and large block reads/writes.

If the workload is predominantly small block I/Os, it is important to examine I/Os per second (IOPS). If the workload is predominantly large block I/Os, it is important to evaluate the throughput in megabytes per second (MB/s).

A good technique for determining if the storage system is servicing predominantly large block versus small block I/Os is to analyze the HP Command View EVAPerf output and divide the total MB/s by the IOPS. This provides the average transfer size. When this result is less than 16 KB, the work load can be considered predominately small block I/Os.

Small block transfers

In the situation where the workload is predominantly small block I/Os, the ability of the storage system to service workloads is primarily bounded by the number of I/Os processed in a given period of time, which is typically measured in IOPS or I/Os per second.

The graph in Figure 7 shows an example of an EVA8000 performing an online controller firmware upgrade with an OLTP work load, with an I/O size of 8 KB. In this example the total time that I/Os are not being process is less than 30 seconds.

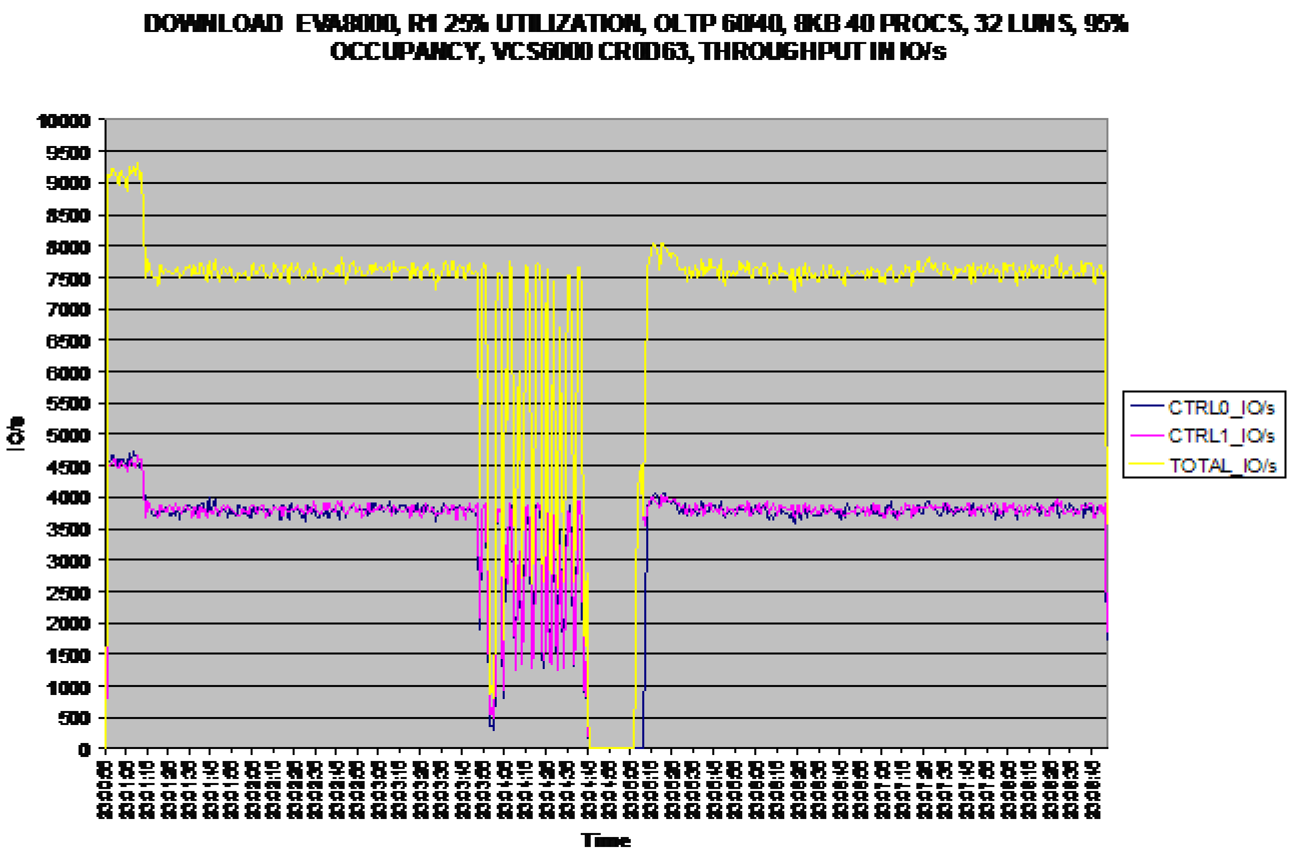

Figure 7 Small block transfer profile

Large block transfers

Unlike small block transfers, when the workload on the storage system is dominated by large block I/Os the workload capacity of the storage system is best monitored in terms of the throughput as measured in megabytes per second or MB/s. Large block sequential workloads are typical of backup and restore operations.

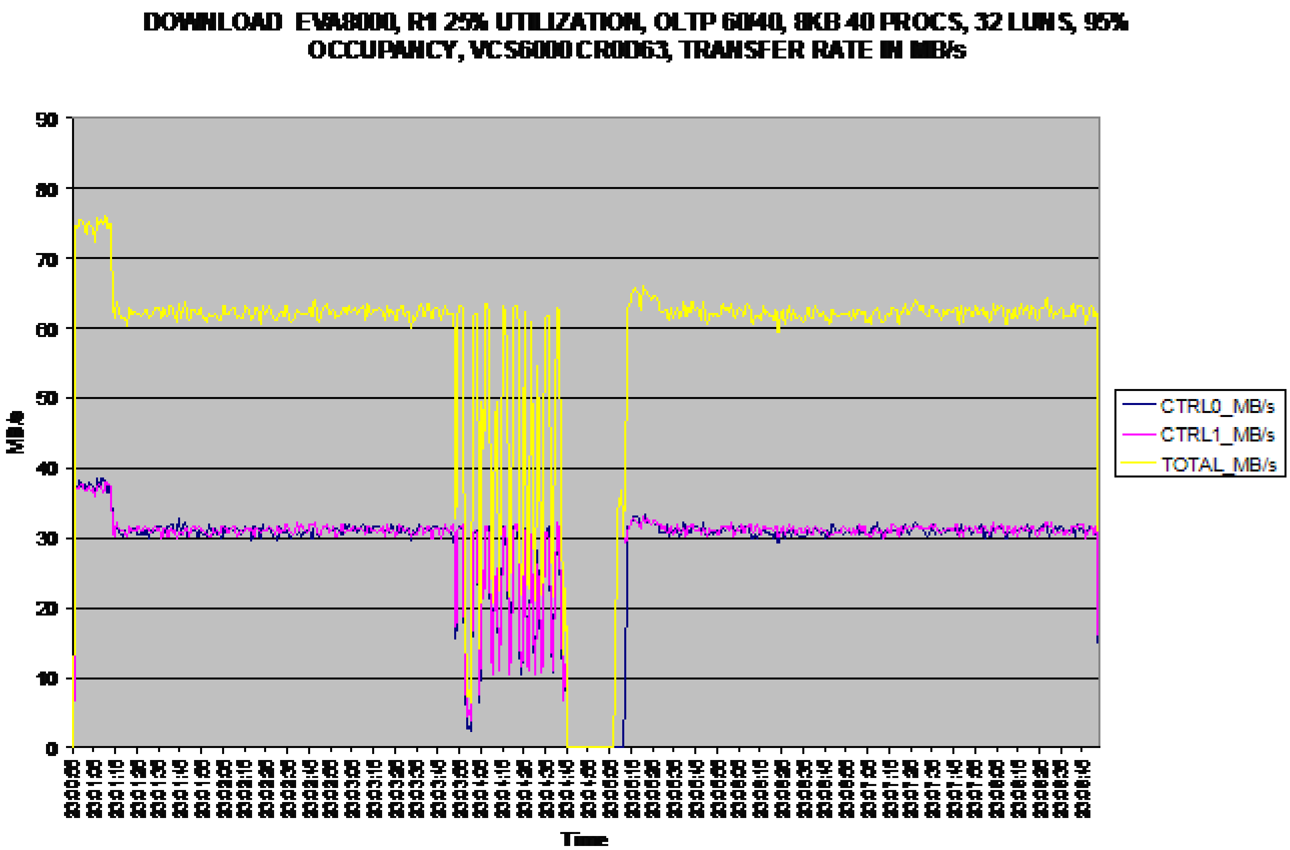

Figure 8 Large block transfer profile

Recommendations for upgrading a busy storage system

In some environments there may not be a period when the storage system load is low enough to safely perform an online firmware upgrade. In this situation there are two alternatives:

- Perform an offline upgrade. Scheduled downtime can be used to perform the upgrade when convenient. This will delay use of the benefits and enhancements included in the new firmware release.
- Use HP Command View EVAPerf statistics to identify very active virtual disks, and then halt only the application(s) imposing the heavy workload. Other applications can remain online during the firmware upgrade. This method has significantly less impact on the SAN than performing an offline upgrade.

  For example, analysis may reveal that a virtual disk is very busy during a particular time because a host is using that virtual disk to perform backups. The identified host's backup application can taken offline, reducing the storage system load to a point where an online upgrades can be safely done.

3 Customer testimonials about online upgrades

A number of EVA customers have successfully used the online firmware upgrade feature. The following are a few of these successes.

Customer 1: A large telecommunications company

Customer 1 is a very large EVA customer with hundreds of storage systems. They use Windows, HPUX and Solaris hosts. Applications include MicroSoft Exchange, MicroSoft Sequel Server, Oracle, and a number of custom applications.

They decided that they needed to upgrade nearly 100 EVA5000 storage systems from VCS 3.020 and 3.025 to VCS 3.028. The magnitude of this upgrade made scheduling of offline upgrades complicated and painful. The customer's management team requested that online upgrades be used to reduce the down time of systems during the upgrade process.

The HP customer support team started worked with engineering to create guidelines and processes that would make this effort successful. Many of those guidelines are included in this document.

The team monitored the storage systems to find appropriate times to do the online controller firmware upgrade. They found that even systems that were supposedly fully utilized generally had times where workloads were at a level where the online upgrades could be done successfully. Once an upgrade window was identified, the upgrade was scheduled. At the time of the upgrade, storage system throughput was monitored to validate that the workload was within the acceptable guidelines. If the demand on the storage system was at an acceptable level, they proceeded with the upgrade.

With nearly 100 storage systems to upgrade there were a few issues. The most significant issue was servers occasionally losing connectivity to the storage system. The customer carefully and methodically investigated the problem. Their investigation revealed that a third-party multitasking driver had been configured with paths that no longer existed. Consequently, when the storage system was resyncing, the driver did a failover to path that was nonexistent. In subsequent upgrades, verification of failover paths was added to the preparation process and the problem was eliminated.

On several storage systems, a time period could not be found where the work load fell off to the desired levels. In these cases, two different approaches were used to upgrade those storage systems.

- Schedule an offline upgrade.
- Identify the host(s) generating the high workloads, and then shutdown just those hosts to reduce the workload to an acceptable level.

Customer 2: A large software company

In the first half of 2006, customer 2 performed at least six online controller firmware upgrades in their data center. Five EVAs were used for internal applications and one EVA was used to host Microsoft Exchange data. All the upgrades were successful.

The most interesting case is the upgrade of the Microsoft Exchange storage system. After careful analysis of the EVA to evaluate the demand on the storage system, the controller firmware upgrade was done around 4 PM during a business day.

Customer 2 also successfully upgraded two EVAs from 2C6D to 2C12D while host I/Os are running.

Customer 3: A large computer chip manufacturer

Customer 3 is another long time EVA customer with EVA storage systems being used at a large number of facilities. Their EVAs host data for a number of systems including Oracle 10g.

Their experience with online upgrades is very different from the others described here. The division that was recently visited had been using online upgrades as their preferred method of controller upgrades for a number of releases. They had been using online upgrades from the early VCS 3.x versions of firmware.

When asked about their experiences, the customer indicated that online controller firmware upgrades work in their environment and were part of their upgrade plans. This customer was one of the first to note the lack of online controller firmware upgrades from XCS 5.031 to 5.100. They were a big part of the reason that the lab developed the XCS 5.032 bridge release.

A Effect of online firmware upgrade on application resiliency

This appendix describes the impact of an online upgrade on various applications.

Testing configuration

The testing consisted of three phases; phase one was used as a baseline, phase two included light application load testing, and phase three included heavy application load testing. The first phase of testing validated the code upgrade procedure without any application load running. Both load tests consisted running the applications under load while repeatedly performing an online code load of firmware on the EVA4000.

The servers used were HP ProLiant servers, with DL380 servers as the application servers and DL360 servers for the load simulation clients, storage management, and OS infrastructure duties. Both load test phases utilized the same server and storage configuration. See Figure 9.

The layout used for the application and databases during testing is shown in Table 4.

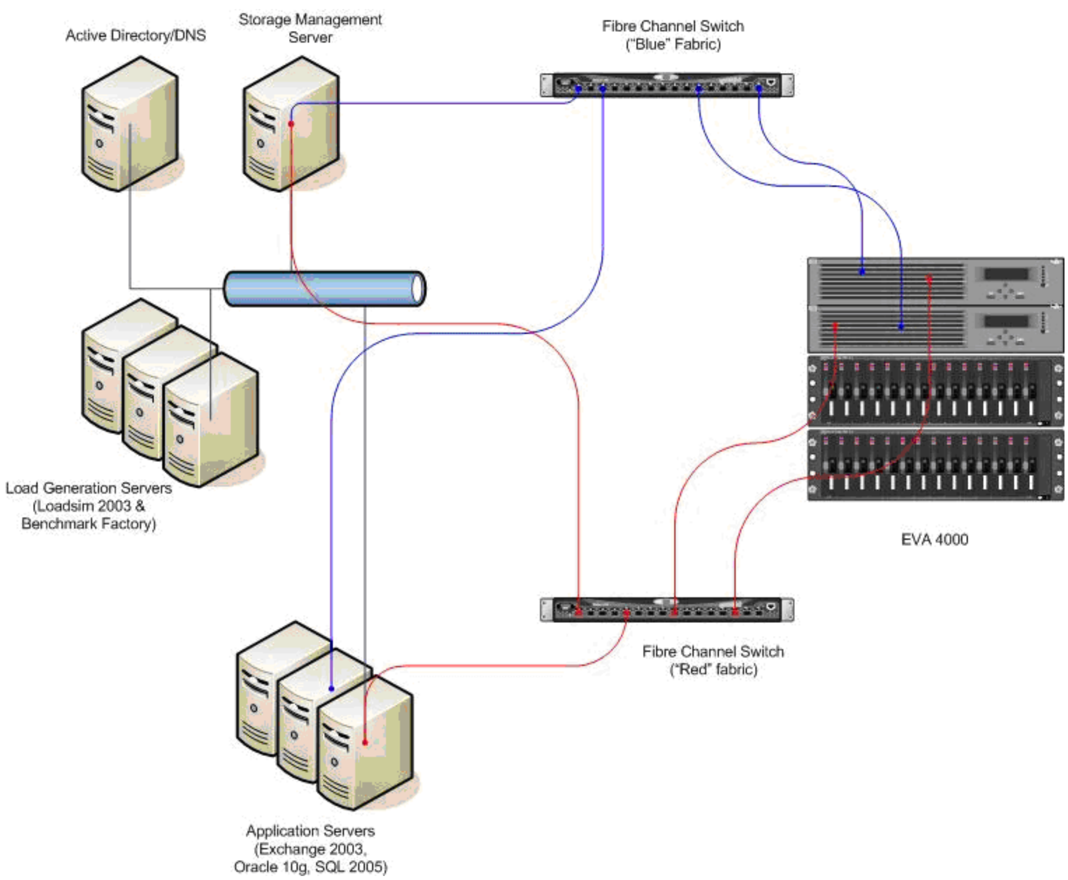

Figure 9 Test configuration

Table 4 Application/Database layout

Microsoft Exchange Server 2003 SP2	
Number of users (light)	750
Number of users (heavy)	2000
Load Profile	MMB3
Number of Storage Groups	2
Number of Databases per Storage Group	2
Mailbox Size	100MB
Load Simulation Tool	LoadSim 2003
Microsoft SQL Server 2005	
Transaction Type	TPC-C OLTP
Read/Write Distribution	70/30
Database Size	120GB
Client Connections (light)	50
Client Connections (heavy)	200
Load Simulation Tool	Benchmark Factory
Oracle 10g	
Transaction Type	TPC-C OLTP
Read/Write Distribution	70/30
Database Size	120GB
Client Connections (light)	50
Client Connections (heavy)	200
Load Simulation Tool	Benchmark Factory

Test cases

Baseline test

Performed baseline test of the online code load operation. The objective was to test the firmware upgrade and determine the length of time the upgrade will take. The expected outcome was a successful non-disruptive upgrade of the EVA firmware.

The test steps included the following:

1. Perform non-disruptive firmware upgrade.
2. Capture firmware upgrade duration.
3. Perform firmware downgrade.

The data to be captured was the duration of firmware upgrade.

Light load testing

Performed multiple online code loads under a light load. The objective was to observe application resiliency to the firmware upgrade. The expected outcome was that the firmware upgrade should complete without the applications failing or having excessive latency.

The test steps included the following:

1. Start load simulators on all three applications.
2. Start capturing metrics with Windows Performance Monitor.
3. Allow I/O to run for at least one hour.
4. Perform firmware upgrade.
5. Save captured metrics.
6. Repeat steps 1-5 two more times.

Heavy load testing

Performed multiple online code loads under a heavy load. The objective was to observe application resiliency to the firmware upgrade. The expected outcome was that the firmware upgrade should complete without the applications failing or having excessive latency.

The test steps included the following:

1. Start load simulators on all three applications.
2. Start capturing metrics with Perfmon.
3. Allow I/O to run for at least one hour.
4. Perform firmware upgrade.
5. Save captured metrics.
6. Repeat steps 1-5 two more times.

Test results

Light load testing

Light load testing revealed that all three applications reported no errors of any kind during the firmware upgrade, while displaying moderate to significant latency for the duration of the firmware upgrade. The firmware upgrade lasted approximately six minutes, during which all three applications continued to respond and did not cause any errors.

While the firmware event did cause latencies that are unacceptable for normal operation, the amount of time the unacceptable values persisted is significantly less than the amount of time that a system shutdown, firmware upgrade and system restart would take.

Microsoft Exchange 2003 performed adequately until the firmware upgrade. See Figure 10. The large spike in LoadSim latency corresponds to the firmware upgrade. Once the upgrade was complete, the Exchange system returned to normal operation with an acceptable latency (95th percentile Loadsim score of < 1000). The average seconds per disk write mirrors the latency spike in the LoadSim score, but also returns to a normal level after the event completes.

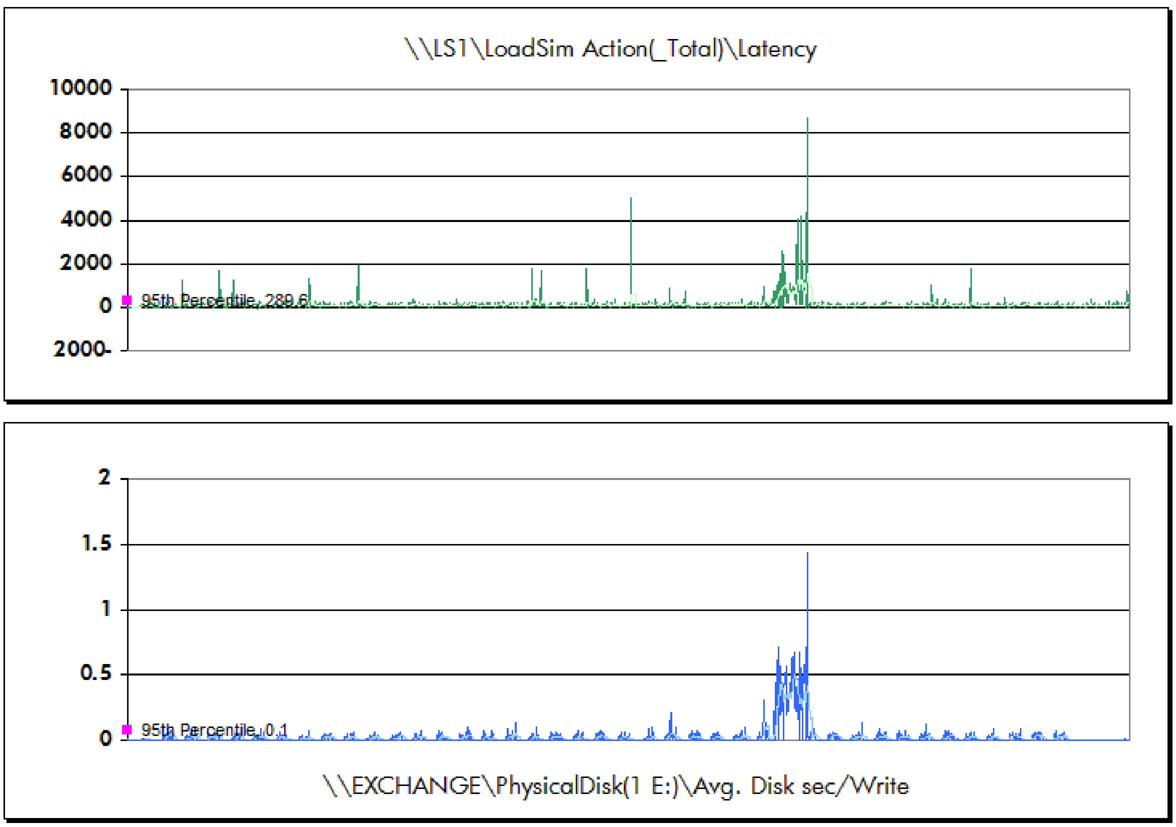

Figure 10 Light load on Exchange 2003

Microsoft SQL 2005 performed well during the firmware event. See Figure 11. While the number of transactions per second did fall during the event, as soon as the firmware upgrade completed, the number of transactions returned to its normal level. The write to disk latency on the LUN holding the SQL 2005 database has a similar increase as the latency on Exchange, but it also recovered quickly to a steady state after the upgrade event.

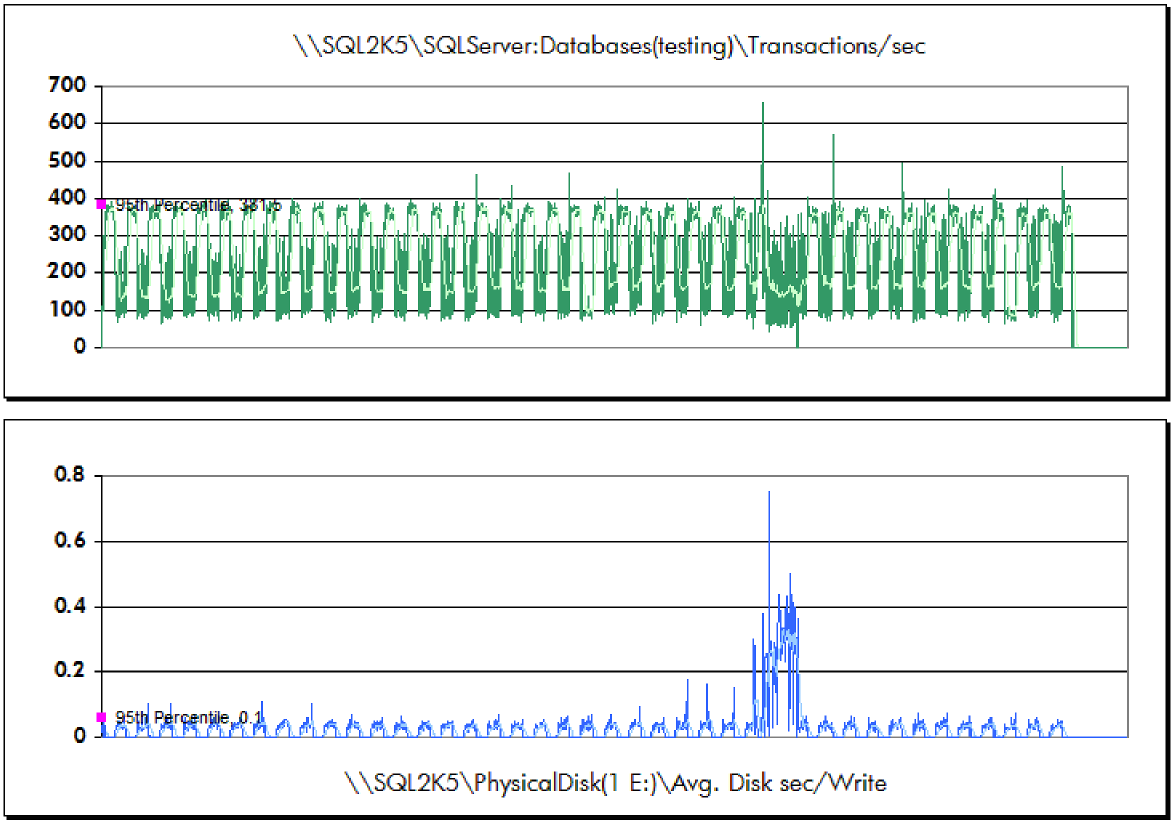

Figure 11 Light load on Microsoft SQL Server 2005

Oracle 10g performed well during the firmware upgrade event. See Figure 12. While during the event, the number of transactions per second dropped off to a fraction of the steady state, the application did not return any errors and ramped back up to its pre-upgrade level.

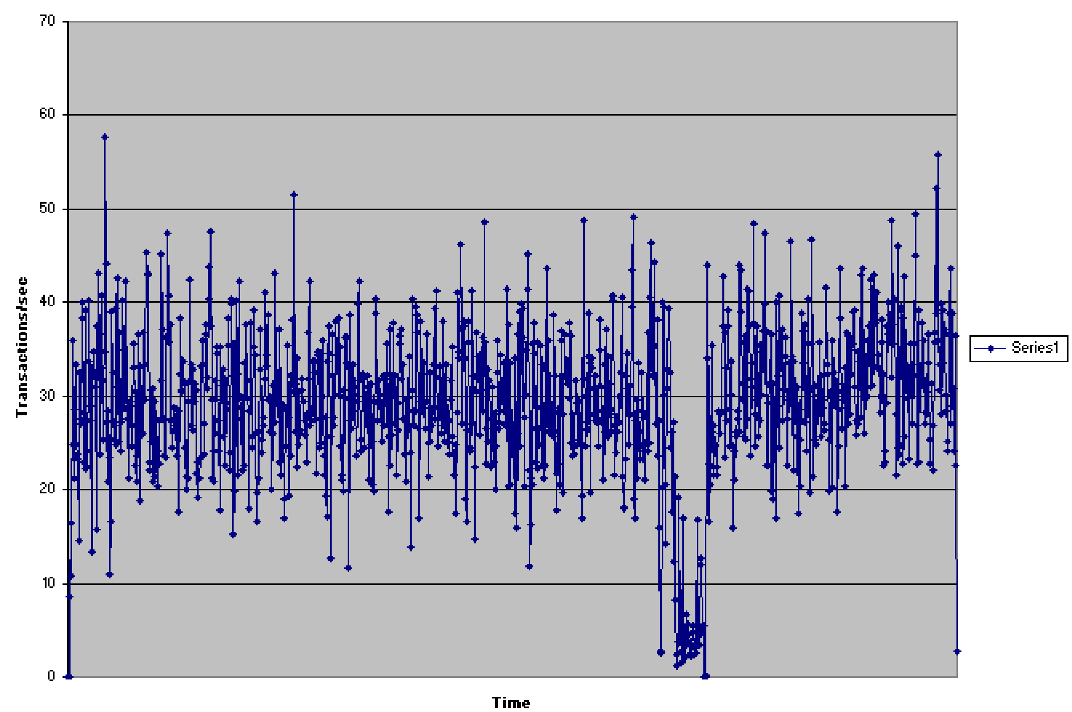

Figure 12 Light load on Oracle 10g

Figure 13 shows the CPU utilization of the two EVA 4000 controllers during the firmware event. The CPU utilization prior to the firmware upgrade is very low (this configuration did not include any HP Business Copy EVA or HP Continuous Access EVA activity). Once the firmware upgrade is started, the CPU utilization spikes to near 100%. Again, once the firmware upgrade is complete, the CPU levels return to normal.

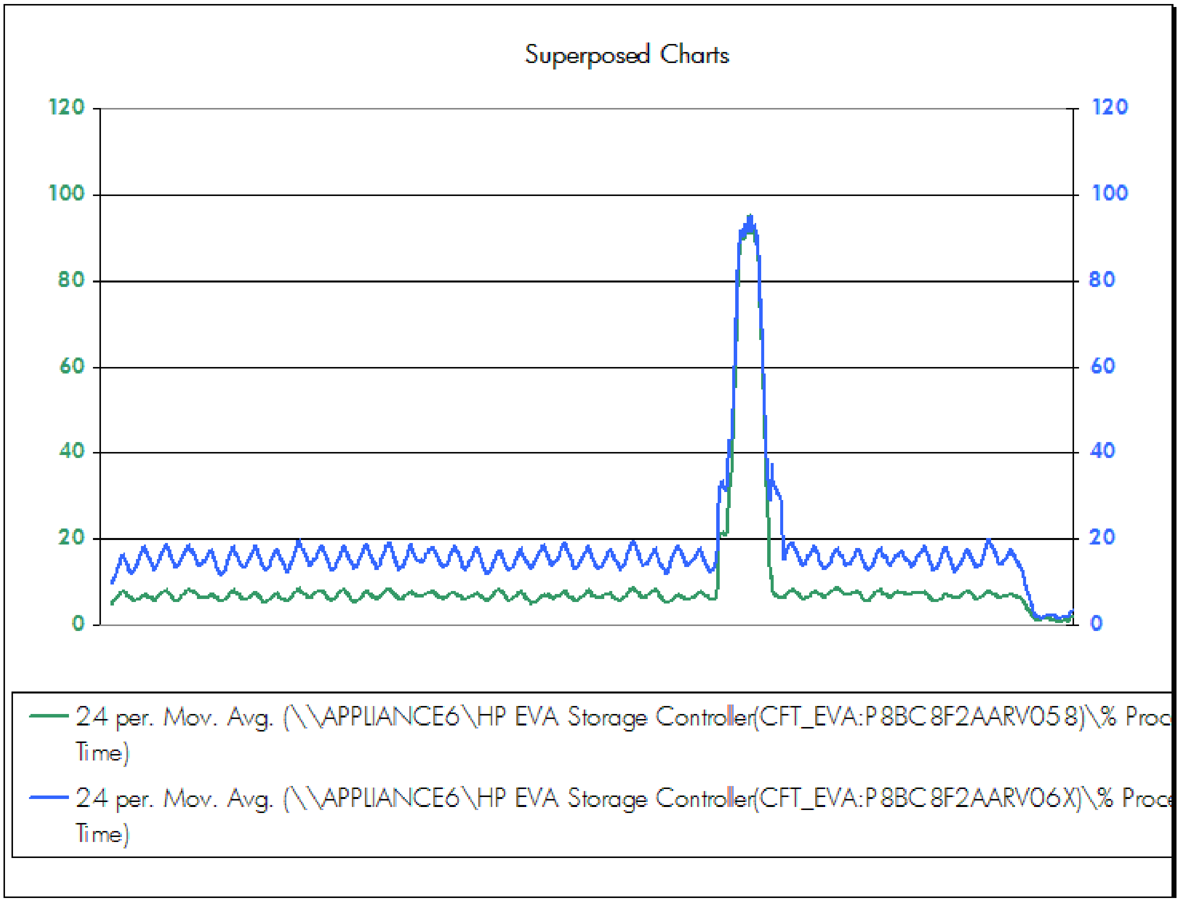

Figure 13 Light load on EVA, CPU utilization

Figure 14 shows the number of I/Os per second on the EVA 4000, achieving a 60-40 split between reads and writes across a representative host port. During the firmware upgrade, I/Os dropped off significantly, but returned to their prior levels immediately after the upgrade event completed.

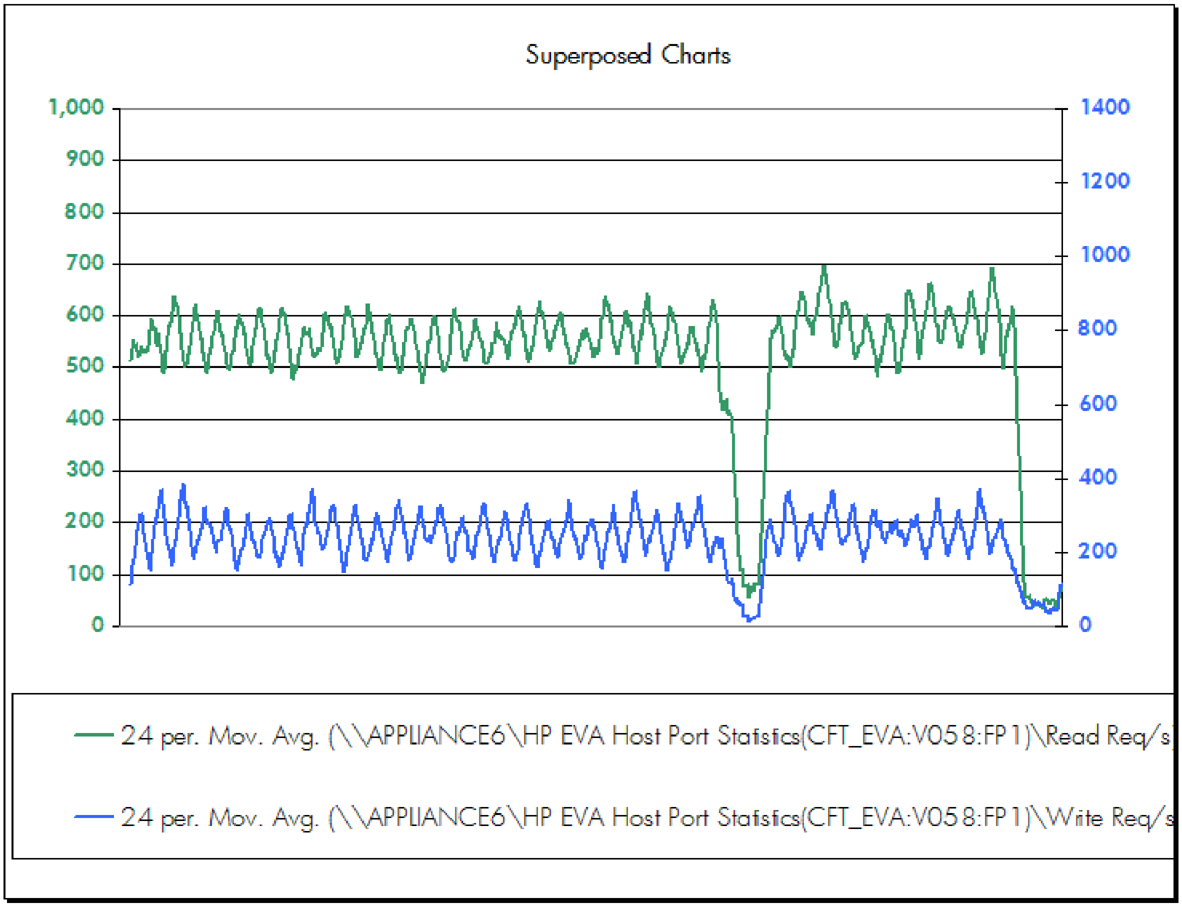

Figure 14 Light load on EVA, I/Os per second

Heavy load testing

Heavy load testing produced very similar results to the light load testing. The latencies and delays during the firmware upgrade event were significantly higher, however. Again, the applications reported no errors of any kind during the firmware upgrade.

During the heavy load phase, two LoadSim client machines were necessary to drive the number of users. The latency score for both servers reached a maximum of ~5000ms, during the firmware upgrade. See Figure 15. This latency translated into queued messages on the Exchange server and in LoadSim. After the firmware upgrade completed, however, the latency dropped back to normal and all queued messages drained. The firmware upgrade lasted approximately six to ten minutes.

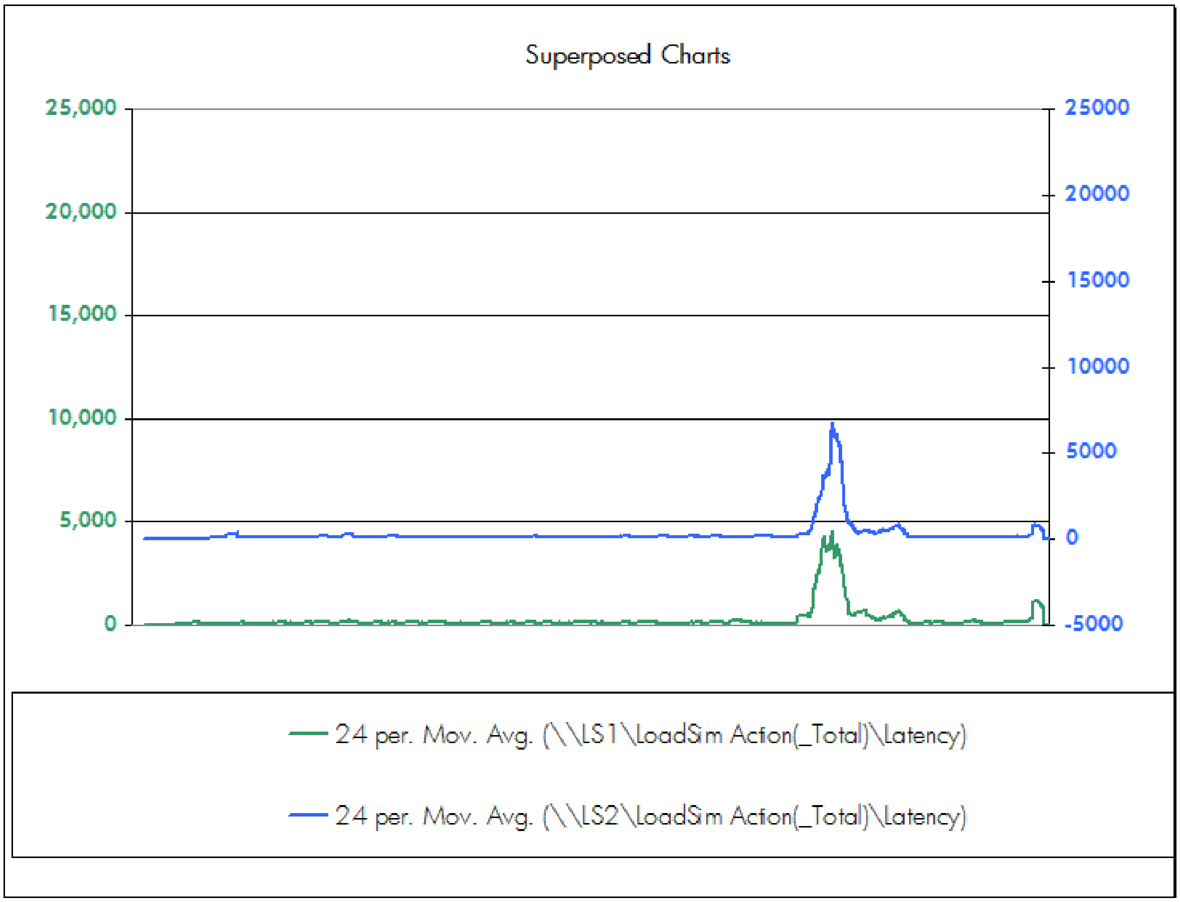

Figure 15 Heavy load on Exchange 2003

Microsoft SQL Server 2005 performed well under the heavy load and during the firmware upgrade. In Figure 16 and Figure 17, the transactions per second are compared against the average disk latency per read and write, respectively. While the disk latencies spiked during the firmware upgrade, the number of transactions per second did not drop to zero. While this does represent a significant performance impact, the application reported no errors of any kind during the firmware upgrade.

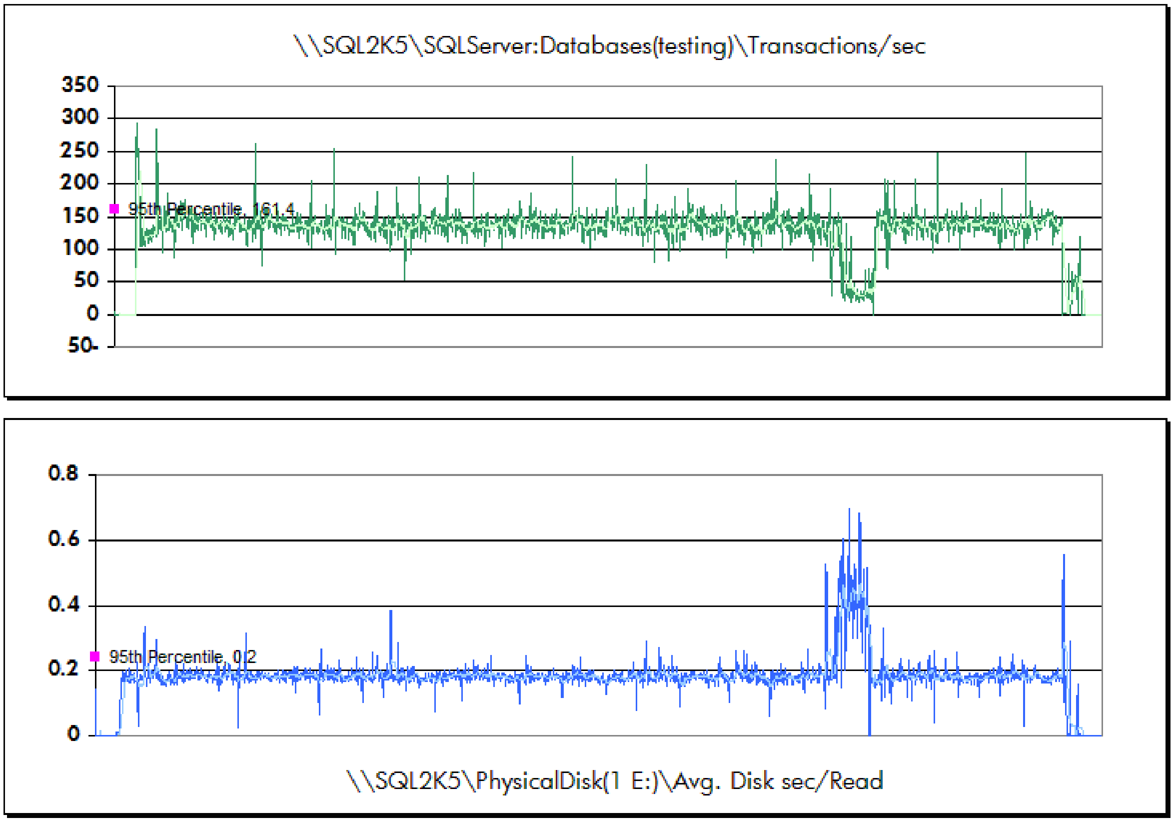

Figure 16 Heavy load on Microsoft SQL Server 2005, sheet 1

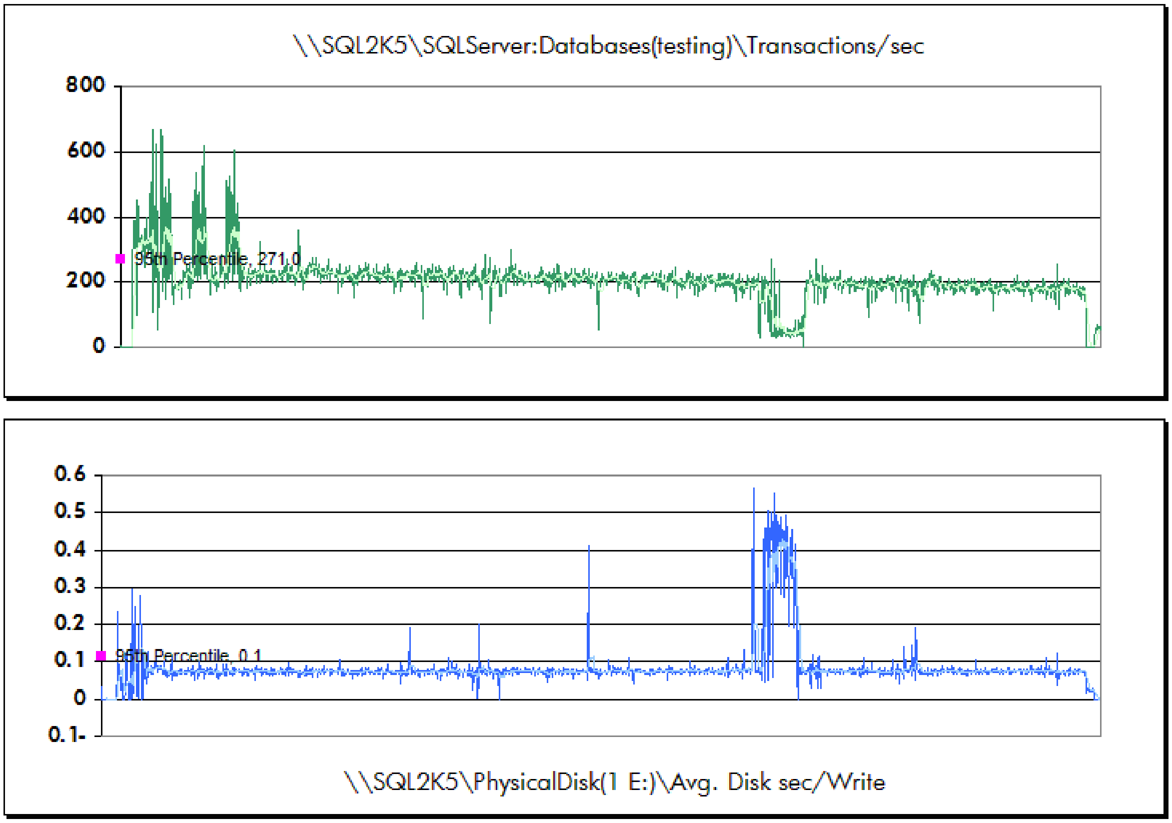

Figure 17 Heavy load on Microsoft SQL Server 2005, sheet 2

Oracle 10g had a very similar performance impact from the firmware upgrade as Microsoft SQL 2005. Figure 18 shows the transactions per second dipping sharply during the firmware upgrade event, but then returning to a steady level once the upgrade is complete. While a reduction of transactions per second to nearly zero is not an ideal situation, the continuous running of the application server is of far more value.

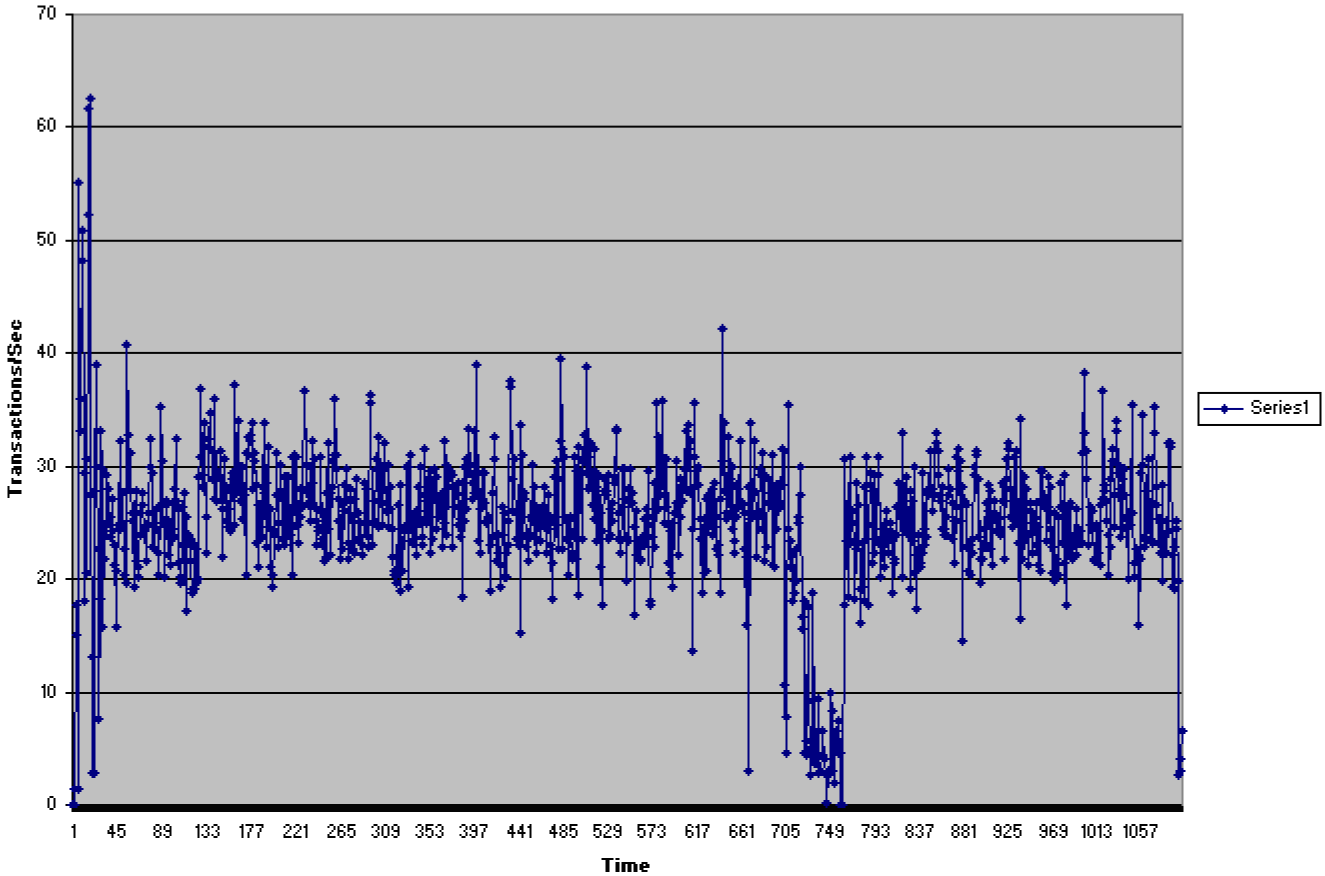

Figure 18 Heavy load on Oracle 10g

The performance of the EVA 4000 is displayed in the following figures, showing the behavior of the array during the upgrade event. First, the CPU utilization spikes for both controllers during the upgrade. See Figure 19. The relatively quick return to normal operating levels means that the applications running on the connected servers are able to weather the event without failure.

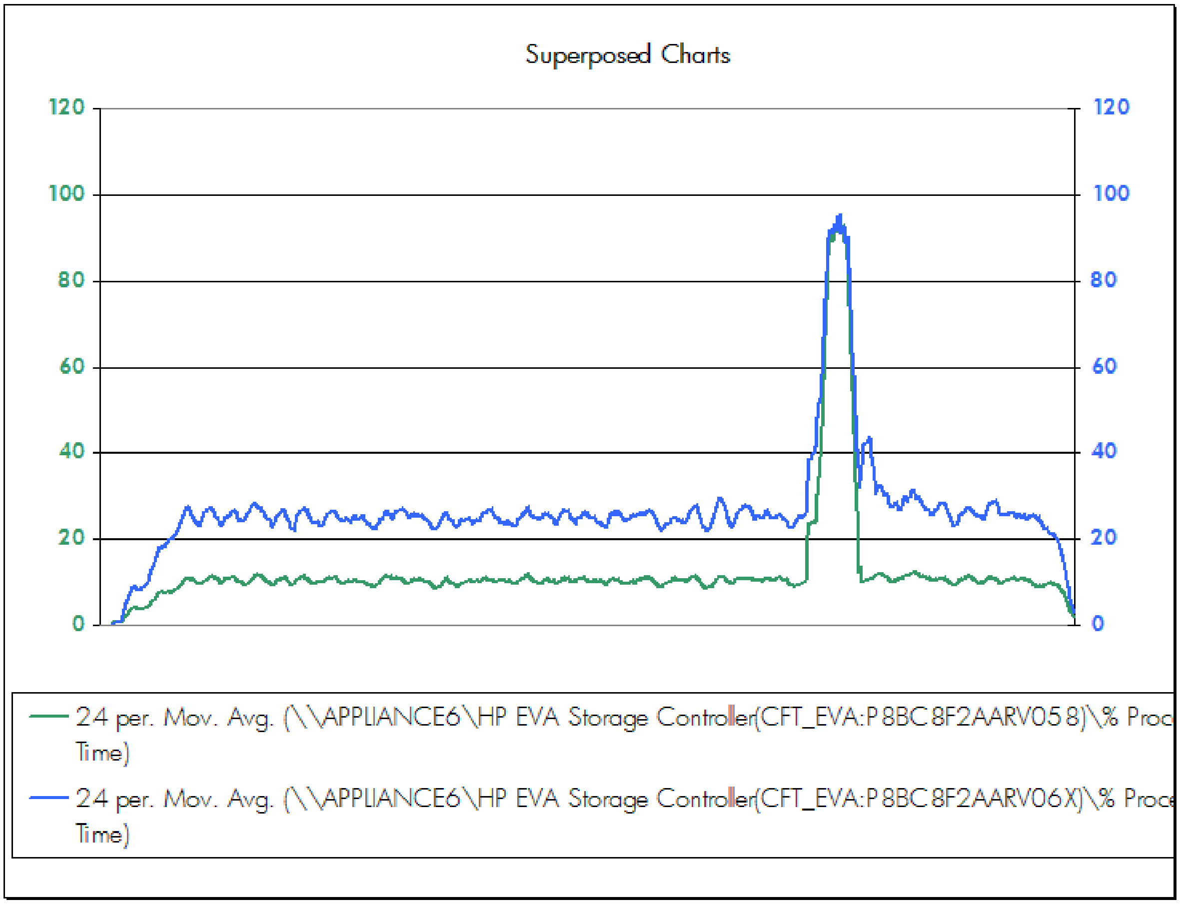

Figure 19 Heavy load on EVA, CPU utilization

Figure 20 shows I/Os per second for both reads and writes. During normal operation, there is roughly a 60-40 split between reads and writes on the port in question. During the firmware upgrade, the number of I/Os per second drops to nearly zero as the EVA 4000 handles the firmware upgrade. They do, however, return to a steady state once the upgrade is completed. Figure 21 shows a combined view of all I/Os per second across the entire array, also displaying the same sharp drop in performance during the upgrade.

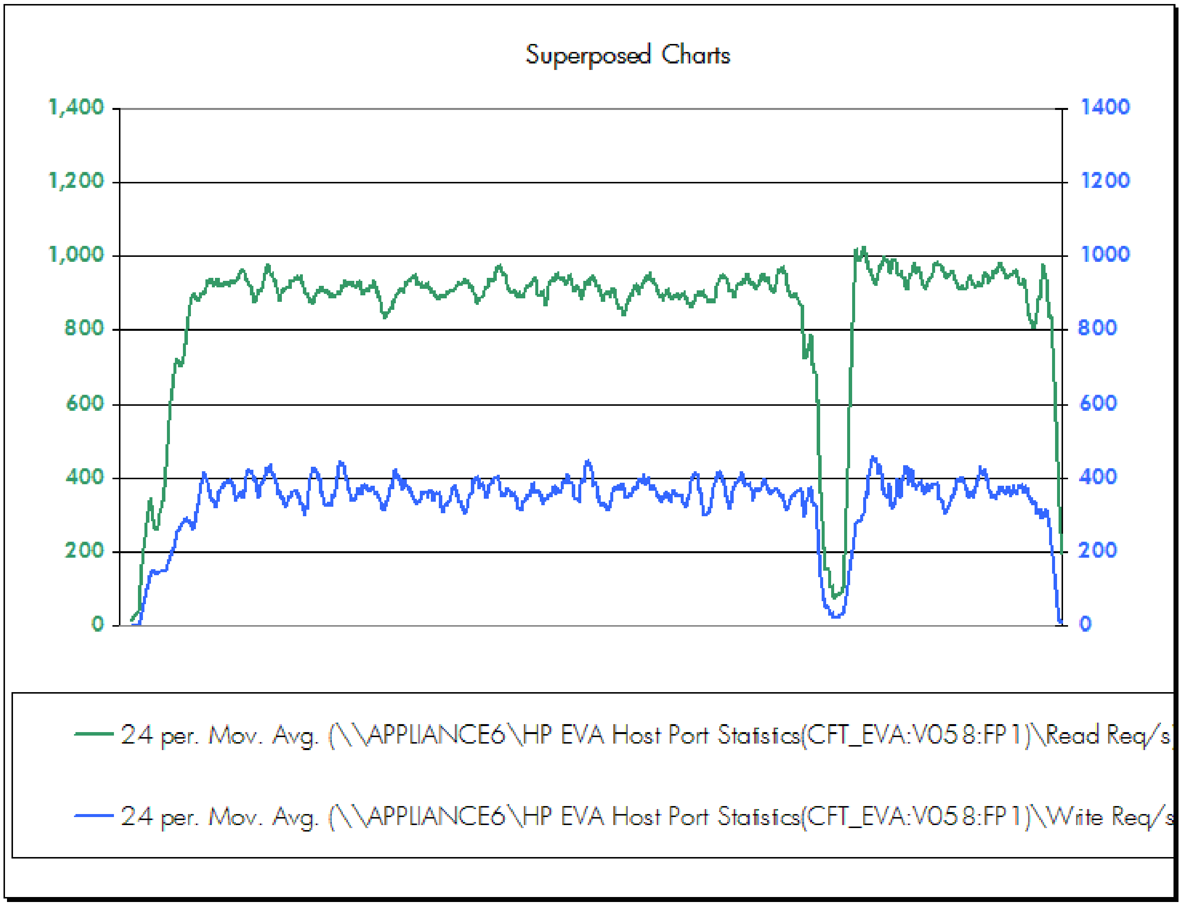

Figure 20 Heavy load on EVA, I/Os per second

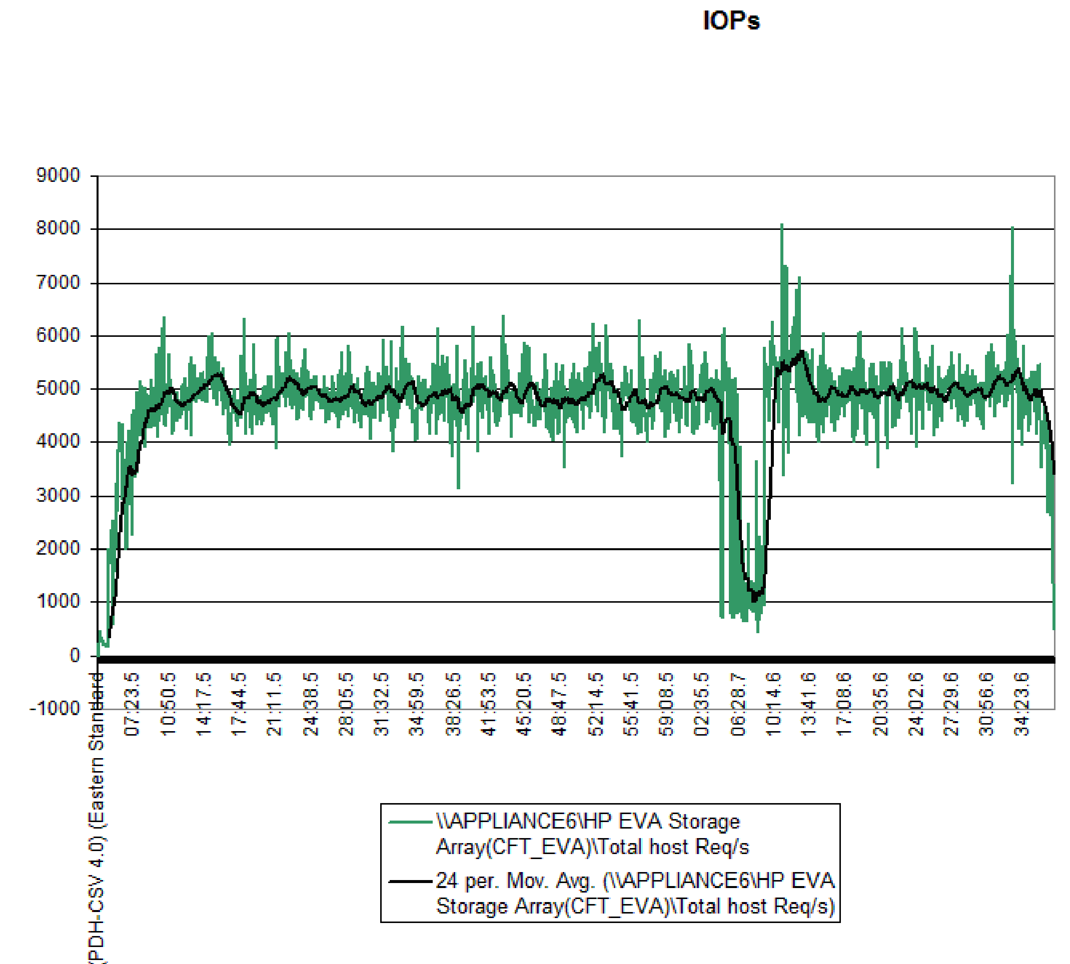

Figure 21 Heavy load on EVA, total I/Os per second

[Figure 22](#) shows the total Kilobytes per second being handled by the array. During the firmware upgrade, the amount of data being handled by the array drops to a mere fraction of what it is capable of during the normal course of operation. When the firmware upgrade is complete, however, the level of data being processed returns to normal. Note that the upgrade event lasts approximately 6 minutes.

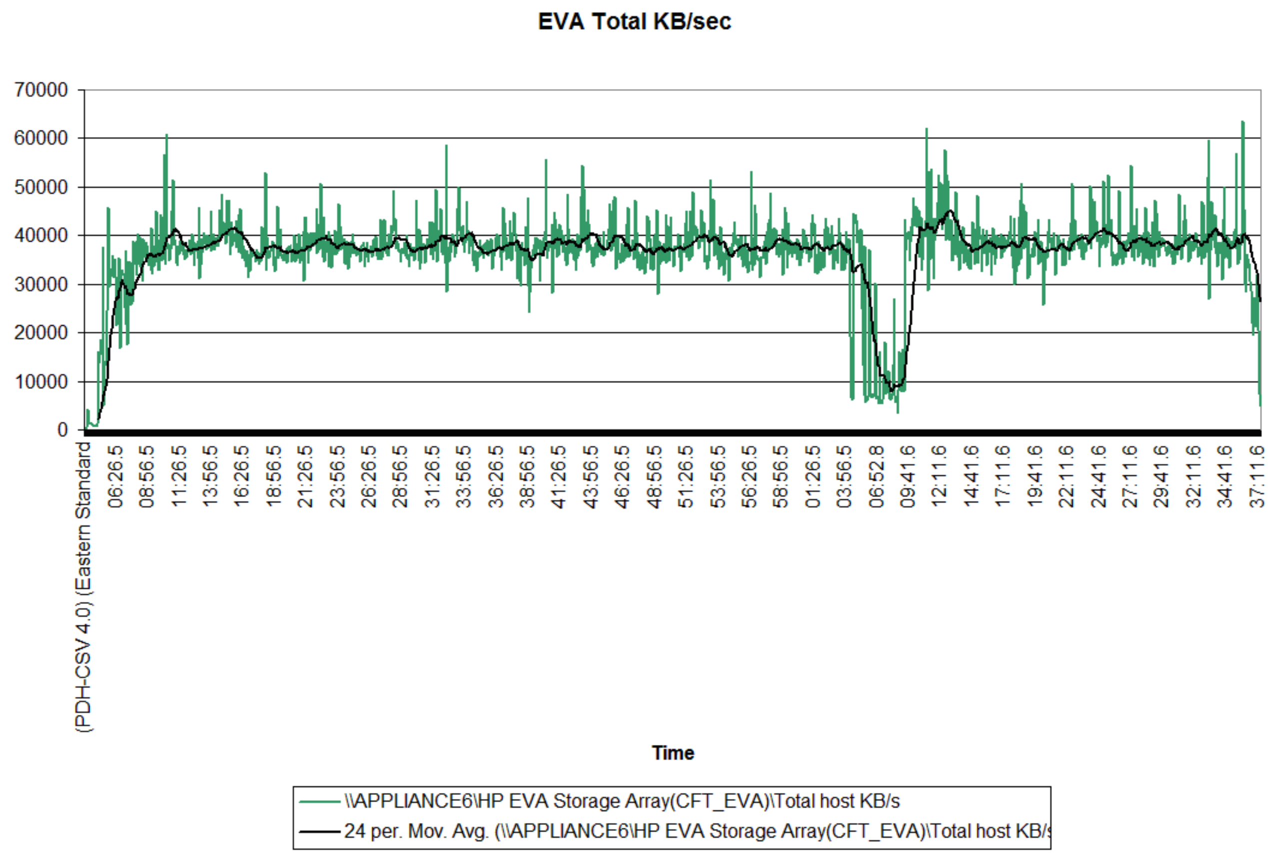

Figure 22 Heavy load on EVA, total KB/s

Conclusion

Two main conclusions can be reached from the data collected by this testing. The first is that the applications tested have sufficient resiliency to report no errors during a normal EVA firmware upgrade, even during load. The programs did not fail during the upgrade event and returned to a steady state immediately after the event was completed. No ramp time was required.

The second conclusion is that there is a definite performance impact to conducting an EVA firmware upgrade during the operation of all three of these applications. The upgrade event itself took between six and ten minutes to complete. During that time, all three applications showed a significant decrease in performance.

From this we can conclude that while the upgrade should complete successfully in the tested environment, the performance impact would be unacceptable during peak business hours and would cause a poor user experience.

Performance metrics

Exchange 2003

- Loadsim 95th Percentile Latency (total) Score
- Read and Write latency (log and database)

Microsoft SQL Server 2005

- Transactions per second
- Seconds per write (log and database)

Oracle 10g

- Transactions per second
- Seconds per write (log and database)

Application Servers

No performance metrics measured.

Storage

- Disk queue depths
- EVA CPU utilization
- Read and write reqs/second
- Total host KB/sec

Index

T

U

V

W