Replicating Microsoft Exchange 2003 using HP StorageWorks Continuous Access and the HP StorageWorks Enterprise Virtual Array white paper

Executive summary

Disaster tolerance to support communications and messaging is an important goal in today's business environment. As communication with email has now become a business-critical activity, providing users immediate and consistent access to Microsoft® Exchange environments is a significant goal. This white paper provides proof points, guidelines, and data points gathered from testing a disaster-recovery solution from HP that meets this objective. The paper provides real-world data that enables appropriate deployment decisions to be made based on specific business requirements. The actual test results provided can help form the foundation for a predictable deployment as well as various recovery scenarios.

This paper describes results from the testing of a replicated Microsoft Exchange Server 2003 cluster configuration accessing Exchange databases stored on an HP StorageWorks Enterprise Virtual Array 5000 (EVA5000) storage subsystem with HP StorageWorks Continuous Access array–based synchronous data replication (DR).

The objectives of this testing included:

- Characterization of application performance in an Exchange Server 2003 Continuous Access environment. Exchange load was simulated using Microsoft Exchange Server LoadSim and Exchange Server Jetstress test tools. This performance testing was carried out using round-trip data replication inter-site link (ISL) latencies from 0 ms to 10 ms over a dedicated 1-Gb Fibre Channel–over–IP (FCIP) link.
- Demonstration of recovery processes of the replicated Exchange Server 2003 cluster environment for different failure scenarios.
- Analysis of the impact caused by Continuous Access write history log merge operations on Exchange performance after the loss and restoration of the ISL.

The key aspects of this solution are that:

- The Exchange environment is duplicated across the local and remote sites.
- The remote Exchange server is in a warm standby mode.
- In the event of a disaster, the remote Exchange servers are brought online after the Exchange Information Stores have been failed over to the remote site.
- After the failover process has completed, all Exchange clients recognize the Exchange server configuration. This configuration provides fast failover times after a disaster.
- Continuous Access provides real-time synchronous data replication between EVA5000s for disaster tolerance.

Key findings

Testing showed that:

- Acceptable Exchange performance is achieved with round-trip data replication latencies between 0 ms and 8 ms.
- A properly configured Exchange environment failed over within 15 minutes in the face of various failures.
- Continuous Access write history log merge operations greatly increased I/O load, thus increasing I/O latencies.

Overview

This paper describes a replicated, disaster-tolerant Exchange server and HP storage solution. This solution consists of Microsoft Exchange Server cluster technology, which provides local site application availability paired with HP StorageWorks EVA Continuous Access data replication for disaster tolerance.

The key aspect of this solution is that the local Exchange server cluster configuration is replicated at the remote site. After a failure, the remote Exchange servers are brought online, thus making the Exchange application available to clients. This solution allows the remote Exchange server cluster to be ready to assume the local Exchange server duties. This, teamed with Continuous Access technology, provides easy and quick failover of the Exchange Information Store before bringing the remote Exchange server cluster online.

This solution has several important advantages:

- Acceptable application performance over varied data replication ISL latencies
- Short failover and recovery times (less than 15 minutes for the tested configuration)
- Ease of failover management

Test environment

Testing was performed on an Exchange Server 2003 cluster environment accessing Exchange storage groups stored in an HP StorageWorks Enterprise Virtual Array 5000 Continuous Access configuration. These tests utilized a duplicate configuration at the local and remote sites. The configuration at each site consisted of the following:

- **Exchange servers**—A three-node active/active/passive Exchange Server 2003 cluster.
- **Servers**—Cluster nodes consisting of HP ProLiant BL20p G2 blade servers running Microsoft Windows® Server 2003 Enterprise Edition and Exchange Server 2003 SP1.
- **Storage**—EVA5000 storage array, EVA 2C112D, with 168 72-GB 10K drives.
- **Domain controllers**—Domain controllers consisting of HP ProLiant DL320 G2 servers running Microsoft Windows Server 2003 Enterprise Edition and configured as active directories and global catalogs.
- **Storage management servers**—HP OpenView General Purpose Servers running on DL320 G2 servers configured with Microsoft Windows Server 2003 Enterprise Edition. HP StorageWorks Command View EVA V3.3 and Continuous Access V1.2 were installed for storage management.
- **Storage area network (SAN) infrastructure**—HP StorageWorks SAN Switch 2/16 running version 3.1.2a of the fabric operating system.

- **ISL**
  - 10-km distance—The 10-km distance used a single spool of 10-km fiber between the two switches, so that no FC-to-IP conversion was required. The latency of light through 10 km of fiber is 0.05 ms.
  - Distances greater than 10 km—These distances were simulated. A 1-Gb IP (Ethernet) dedicated link between two HP StorageWorks IP Storage Router 2122-2s was used. A single Empirix Packet Sphere Model 200 delay simulator between the two IP Storage Routers was used to induce round-trip inter-site latencies between 0 ms and 10 ms.
- **Application load simulation**—The LoadSim client configuration consisted of eight HP ProLiant DL320 servers running Microsoft Windows XP Professional SP2, Microsoft Office Outlook 2003, and Microsoft Exchange Server Load Simulator 2003.

The test configuration is shown in Figure 1. For more information, see "Appendix A—Configuration bill of materials (BOM)."

Figure 1. Continuous Access Exchange solution configuration
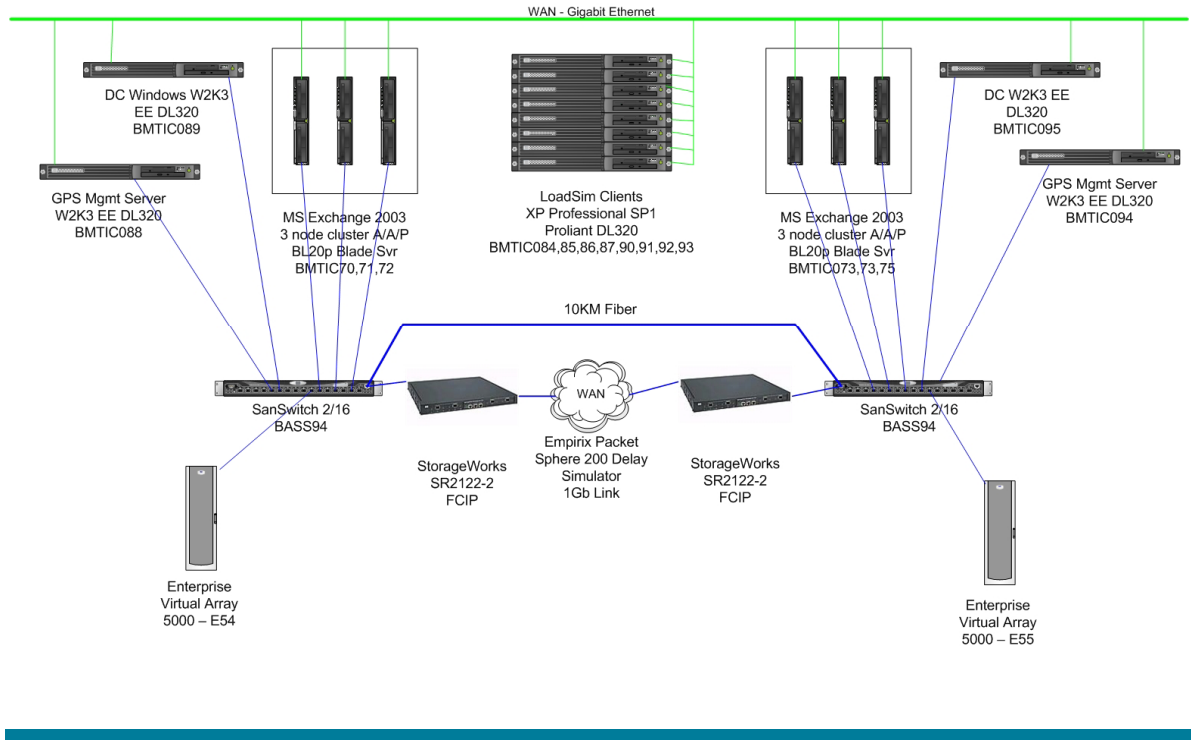
Exchange storage configuration

The Exchange Information Store was configured as follows:

- Phases 1 and 3—Four storage groups per Exchange server, five databases per storage group, and 100-MB mailbox quotas for each user.
- Phase 2—Two storage groups per Exchange server, five databases per storage group, and 100-MB mailbox quotas for each user.

Data replication was configured using one DR group per storage group, where each DR group contained two copy sets: database disk and log disk. For details on storage configuration, see "Appendix B—Exchange storage configuration (Phase 1)."

Phases of testing are described in "LoadSim performance testing" on page 10.

IP configuration

The local and remote Exchange server clusters resided in the same IP subnet. Procedures for using different IP subnets at the local and remote sites were not tested but have been implemented.

Best practices

Best practices for setting up Exchange server cluster

Recommended best practices for configuring an Exchange server cluster environment were used in setting up the test environment. The following documents provide a comprehensive list of best practices (see "Appendix D—Reference documentation" on page 62 for web links):

- *HP Best Practices for Microsoft Exchange Server 2000 and 2003 Cluster Deployments*
- *Exchange Server 2003 Clusters on Windows Server 2003*

Best practices for setting up Exchange in a Continuous Access data replication environment

In addition, the following recommended best practices for configuring Microsoft Exchange in a Continuous Access data replicated environment were taken from the following Microsoft and HP KB articles (see "Appendix D—Reference documentation" on page 62 for web links):

- "Deployment Guidelines for Exchange Server Multi-Site Data Replication," Microsoft Corporation
- "Multi-site data replication support for Exchange 2003 and Exchange 2000," Microsoft Corporation

Best practice recommendations are:

- Use four storage groups per Exchange virtual server to maximize replication performance. Maximizing storage groups and databases within each storage group spreads the I/O load across more logical volumes.
- Minimize the number of users per Exchange virtual server instead of consolidating for replication. Consolidating users increases performance degradation.
- Test on the same network (for latency) and hardware on which production is planned.
- Design for Continuous Access DR up front to avoid implementing on untested configurations. Adding replication later could require major redesign of your Exchange configuration.
- Over-provision the inter-site network to reduce write latencies due to queuing. Plan extra bandwidth and redundant ISLs. If one path fails, the remaining path must handle the load of both paths.
- Increase the transaction log buffer size to reduce the frequency of capacity flushes and increase the log write size, which will reduce the overall log write latency.

Best practices for replicating data

Required replicated data

Database files store message data. There are two types of databases: private mailbox and public folder. Each database consists of two files: the database file (.edb), which holds messages and MAPI content, and the native content "streaming" file (.stm), which has non-MAPI data. Log files (.log) record each transaction to the database. Checkpoint files (.chk) contain the information on the entries in the log files that have been written to the disk. All of these files are vital to provide client access to the mailbox and/or public folder server. Therefore, all of this data is required to be replicated in a synchronous environment.

Recommended replicated data

SMTP and MTA local queue data is temporarily stored while being processed, to prevent data loss in case of server failures. For example, when the destination server is unreachable, the messages to be routed to that server are stored on the local server queue directory. If the disk that stores the queue data fails, all the messages in the queue would be lost.

Optional replicated data

The message-tracking log contains the information on all the messages transferred to and from the server. This data is important for diagnostic purposes. Message tracking is not enabled by default, but if this data is important to the customer's business, it could be replicated.

Microsoft Exchange Server Best Practices Analyzer tool

Microsoft provides the Microsoft Exchange Server Best Practices Analyzer tool (EXBPA) that will examine an Exchange Server deployment and determine if its configuration conforms to Microsoft's best practices. The EXBPA creates a report showing any nonconformance to best practices. See "Appendix D—Reference documentation" on page 62 for a website pointer to EXPBA.

Performance testing

Performance was tested using two different test utilities: Microsoft Exchange Server Load Simulator (LoadSim) and Microsoft Exchange Server Jetstress Tool. Each utility allows Exchange administrators the ability to test their configuration in different scenarios to ensure that their configuration performs properly before going into production.

LoadSim is a benchmarking tool that allows Exchange administrators to simulate Exchange mail message traffic between Exchange clients and Exchange servers. LoadSim executes on client servers, each simulating some number of clients and generating a specific mail load to the Exchange servers. LoadSim gathers client statistics during each test run. The Performance Monitor utility is used to gather further data about each client and Exchange virtual server.

Jetstress verifies the performance and stability of the Exchange disk subsystem. Jetstress is executed on the systems where the Exchange virtual severs are to reside. Jetstress generates the I/O load of a specified number of users by simulating that load on the database and transaction log disks. During the test run, Jetstress gathers performance data for further analysis.

Data replication inter-site latency testing

Performance data was gathered by executing performance tests using different ISL latencies between the local and remote sites. These round-trip latencies varied from 0 ms to 10 ms.

The Empirix delay simulator uses a latency value to create a delay that simulates packets traveling between sites. In actual customer situations, inter-site latencies vary greatly. The physical distance between sites is not a good indicator of latency. For example, a customer has a local site in Colorado Springs and a remote site in Denver. The physical distance is approximately 80 km (50 miles). The latency of light traveling through 80 km of cable is 0.4 ms. Therefore, round-trip data replication latency is 0.8 ms. Latency speeds like this are rarely the case. In most instances the latency is much greater, due to the actual path the data travels and any gateways or routers that process each packet. Packet processing time by each router and/or gateway adds to the latency.

> **Note:** Local service providers should be contacted for actual latencies between sites.

For calculating latencies and the impact of reduced bandwidth, see recommendations in:

- *HP StorageWorks Continuous Access Planning Guide*—AA-RW1DA-TE
- *HP StorageWorks Continuous Access EVA user guide performance estimator*—AA-RU5ME-TE

For pointers to these documents, see "Appendix D—Reference documentation" on page 62.

> **Note:** All latency values used in this document refer to round-trip inter-site latencies (RTIL) between the local and remote sites. The latencies specified do not include the additional latency added by the FCIP gateways and the Empirix delay simulator in the tested configuration.

Baseline testing

A set of baseline tests were performed to form a basis for testing. Two baseline tests were performed:

- **No-replication test**—LoadSim and Jetstress were executed against an Exchange Information Store that was not being replicated. This provides a reference for the additional latency of data replication and increased inter-site latency.
- **0-ms DR latency test**—LoadSim tests were executed with a 0-ms delay. This test used the ISL through the FCIP gateways and the Empirix delay simulator with a zero delay. The test provided measurement of the latencies induced by the FCIP gateways. Figure 2 shows that the latency of the FCIP gateways and the delay simulator, when compared with the no-replication data, added 1.125 ms to the round-trip DR latency of an average log disk write.

Figure 2. Effect of FCIP gateways on latency

Added latency of FCIP gateways and delay simulator

	No replication	10-km fiber (0.1-ms)	0-ms gateways	1-ms gateways
■ Avg DB sec/write	4.375	12.625	28.875	36.25
■ Avg log sec/write	0.875	1	2	3.125

DR Latency (round trip)
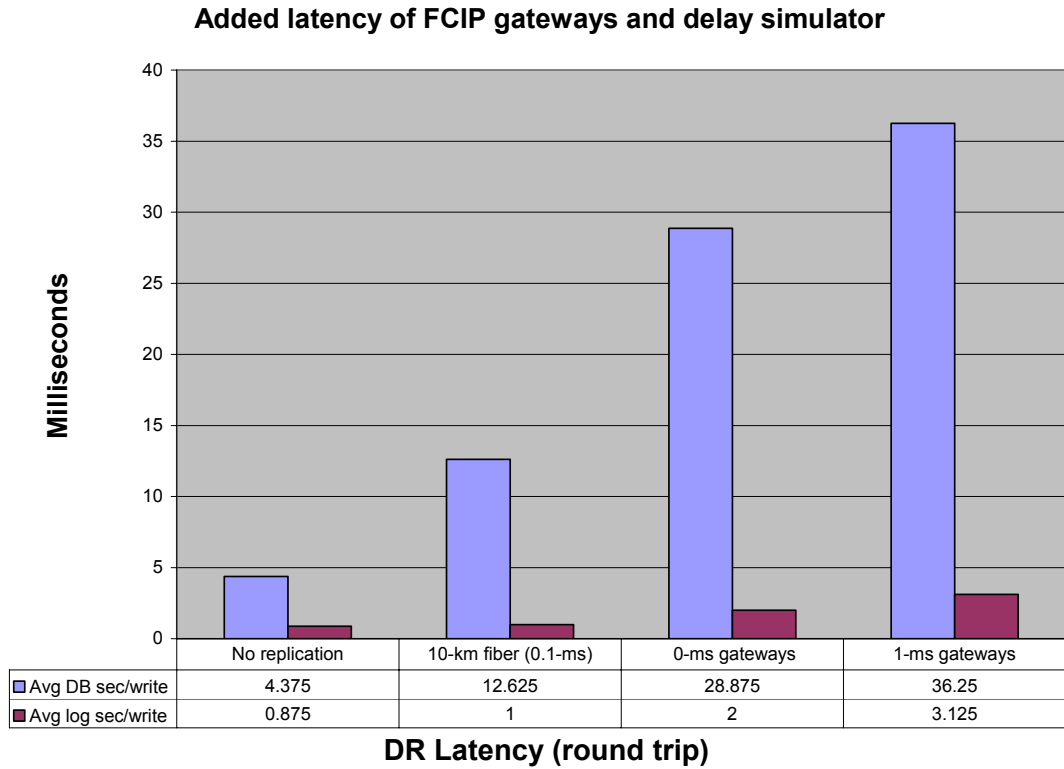
The performance data presented in this document includes these baselines. Use this data for comparison with the DR latency data to see the real effect that data replication, inter-site latencies, and hardware-induced latencies have on performance.

LoadSim performance testing

Performance testing was conducted in three phases, with each phase having different goals:

Phase 1—Execute LoadSim testing using a 5,000-user load over eight storage groups. The number of storage groups was maximized to spread I/O load over more logical volumes. Inter-site latencies ranged from 0 ms to 10 ms. The goal was to see how DR latency affected Exchange performance.

Phase 2—Execute LoadSim testing using a 5,000-user load over four storage groups to double the user load on each storage group and reduce the number of logical volumes. Inter-site latencies ranged from 0 ms to 10 ms. The goal was to see how increasing the user load on storage groups affected Exchange performance.

Phase 3—Execute LoadSim testing by increasing the user load to see the effect on the tested Exchange configuration at a 1.0-ms inter-site DR latency. The user loads tested were: 5,000, 7,500, and 10,000 users. Phase 3 testing used eight storage groups.

Test Phase	Phase 1	Phase 2	Phase 3
Number of users	5,000	5,000	5,000, 7,500, and 10,000
Number of storage groups	8	4	8
Inter-site DR latency	0 to 10.0 ms	0 to 10.0 ms	1.0 ms

Microsoft-recommended performance values for Exchange

Microsoft, through their testing, has recommended performance limits to ensure a positive customer experience. Of importance are disk read and write latencies, Exchange server request (RPC) latencies, and client response times. Microsoft recommendations, from "Deployment Guidelines for Exchange Server Multi-Site Data Replication," are as follows:

- Database disk read average latencies are less than 20 ms, with maximum values less than 40 ms.
- Log disk write and read average latencies should remain less than 20 ms, with maximum values less than 40 ms.
- Exchange server request (RPC) average latencies should remain less than 50 ms, with maximum values less than 100 ms.
- Client 95th percentile weighted response times should remain less than 1000 ms.

Performance values in excess of these recommendations can result in less than ideal customer experience.

Phase 1 test results summary

Phase 1 LoadSim testing showed that Exchange performance values remained within recommended limits except for database write latency when the round-trip inter-site DR latency remained at or below approximately 8 ms. At this 8-ms latency:

- Log write average latencies were 20 ms.
- Log read latencies were well below 20 ms, at about 7 ms.
- Database write average latencies were 130 ms. These higher database write latencies may have had less effect, since database writes typically occur some time later than log writes when database pages are flushed from the cache.

- Database read average latencies were below 20 ms, at about 16 ms.
- Exchange server request (RPC) latencies were 36 ms.
- Client 95th percentile weighted response was just above 1000 ms at about 1050 ms.
- Message throughput began to drop past the 7.5-ms DR latency.

From a business perspective, these results demonstrate acceptable Exchange performance below an 8-ms inter-site latency while achieving the goal of disaster tolerance. Data shows that DR latencies longer than 8 ms begin to affect Exchange performance. Exchange performance and throughput begins to degrade at higher latencies.

Phase 2 test results summary

Phase 2 testing results showed that with an increased load on each storage group and reducing the number of logical volumes caused Exchange performance to degrade at a faster rate than in the Phase 1 testing. Phase 2 testing showed that performance was acceptable to a 7-ms inter-site latency versus the Phase 1 8-ms DR latency.

Test results were:

- Log write average latencies were 20 ms.
- Log read latencies were well below 20 ms at approximately 8 ms.
- Database write average latencies were 130 ms. These higher database write latencies may have had less of an effect since database writes typically occur later than log writes when database pages are flushed from the cache.
- Database read average latencies were at the 20-ms limit.
- Exchange server request (RPC) latencies were 38 ms.
- Client 95th percentile weighted response was 1100 ms.
- Message throughput remained consistent.

Phase 2 testing shows that Microsoft recommendations of maximizing the number of storage groups and databases to increase the I/O load across more logical volumes does result in a performance increase, allowing additional latency to be added for data replication.

It is interesting to note that when Microsoft-recommended limits are crossed, the performance values are very similar for Phases 1 and 2. Both phases degraded in a similar fashion.

Phase 3 test results summary

Phase 3 testing showed that in the tested configuration, performance began to degrade below Microsoft-recommended performance limits at user loads higher than 7,500 users. The data did not show all performance values crossing the limits simultaneously. However, Microsoft message queues did begin to backup with the 7,500-user load. At the 10,000-user load, the many Exchange message queues were extremely long.

LoadSim configuration

Exchange mail message traffic was generated using eight LoadSim clients simulating 5,000 users running the MAPI Messaging Benchmark 3 (MMB3) workload. LoadSim was configured differently for Phase 1 and Phase 2.

For Phase 1, each client server, simulating 625 Exchange users, accessed a single storage group. LoadSim was configured as shown in

Figure 3 and Figure 4.

Phase 2 testing also simulated 625 users per client, but there were two clients per storage group generating a load of 1,250 users per storage group.

The MMB3 workload was designed to model a typical corporate email environment. For more information on the MMB3 workload, see: http://www.microsoft.com/exchange/evaluation/performance/mmb3.mspx

Each LoadSim test run lasted eight hours. The data shown was taken from the steady-state middle four hours of each test period. During LoadSim test runs, performance data was collected by LoadSim and the Performance Monitor tool.

Figure 3. Phases 1 and 3 LoadSim test topology data
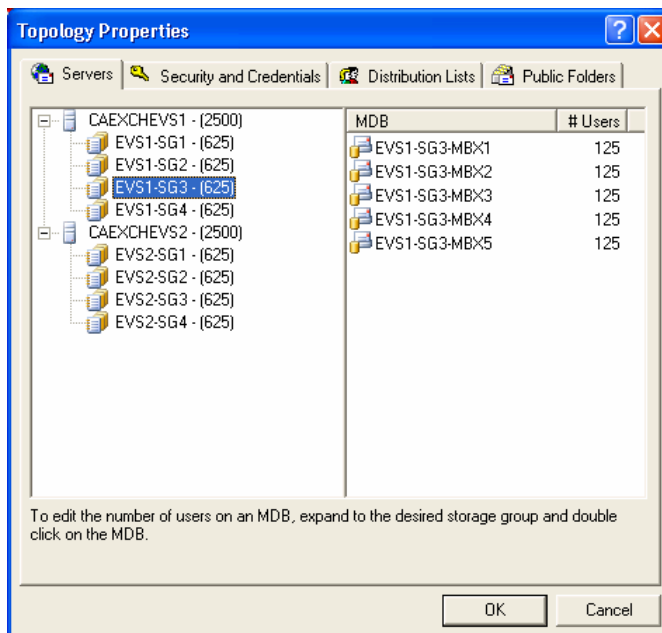
Figure 4. Phases 1 and 3 LoadSim test property data
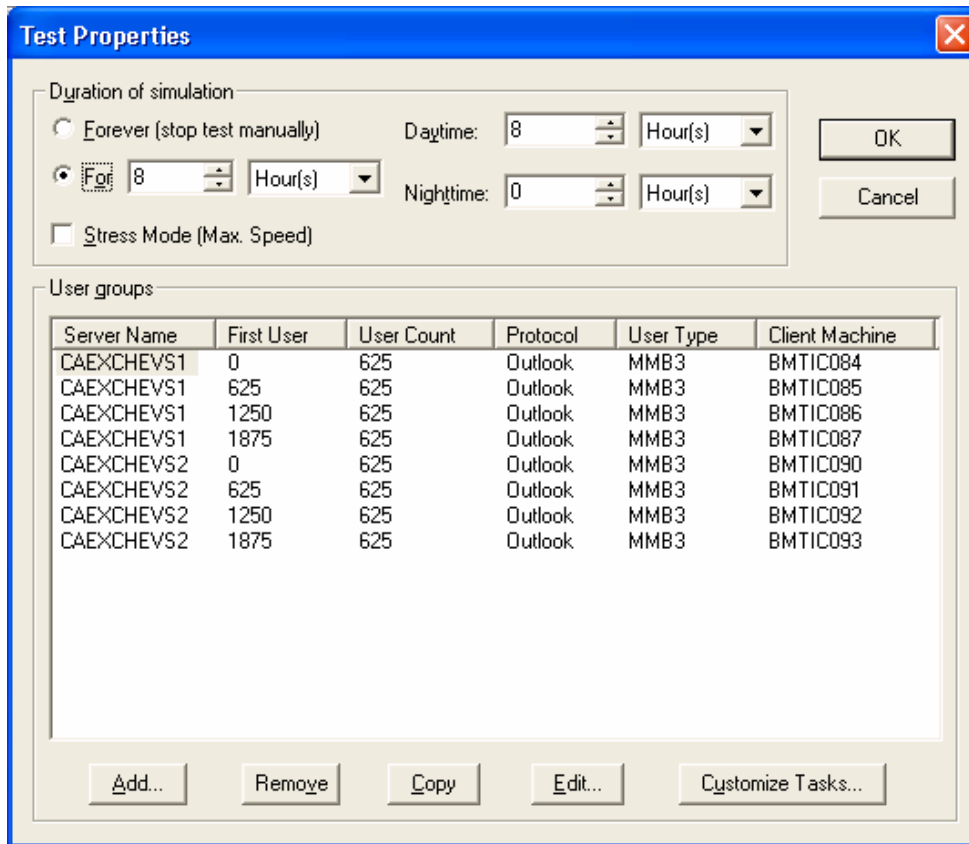
Phases 1 and 2 LoadSim client performance results

LoadSim results indicate that client performance remained below recommended upper limits as the inter-site latency was increased. LoadSim client performance was measured using:

- LoadSim 95th percentile scores
- Client CPU usage
- LoadSim client latency performance values

LoadSim 95% scores, which appear in Figure 5, show the average weighted response times for all Exchange clients at the 95th percentile. This score indicates an overall value based on many Exchange operations, such as send, read, reply, delete, and forward. The value represents a weighted response time in milliseconds. Test results show the 1000-ms limit was crossed at 7 ms for Phase 2 and 8 ms for Phase 3.

The LoadSim 95% score represents what a mail user might experience with the MMB3 workload. Figure 5 demonstrates that a user would still see reasonable response times over the range of mail operations which that user performed during a given day. The score indicates that, 95 out of 100 times, a user would experience acceptable response times with inter-site latencies less than 7 ms and 8 ms, for Phases 1 and 2, respectively.

> **Note:** Test results for Phases 1 and 2 are shown together for comparison purposes in the graphs in the following figures. Phase 1 is denoted by "8 Storage Groups" or "8 SG" for testing against eight storage groups. Phase 2 test results are denoted by "4 Storage Groups" or "4 SG" for testing against four storage groups.

Figure 5. LoadSim 95% score

5000 user LoadSim 95% score

	0-ms gateways	2.5-ms gateways	5.0-ms gateways	7.5-ms gateways	10.0-ms gateways
8 storage groups - < 1000	344	419	539	892	1708
4 storage groups - < 1000	406	509	707	1191	1831

DR latency (round trip)
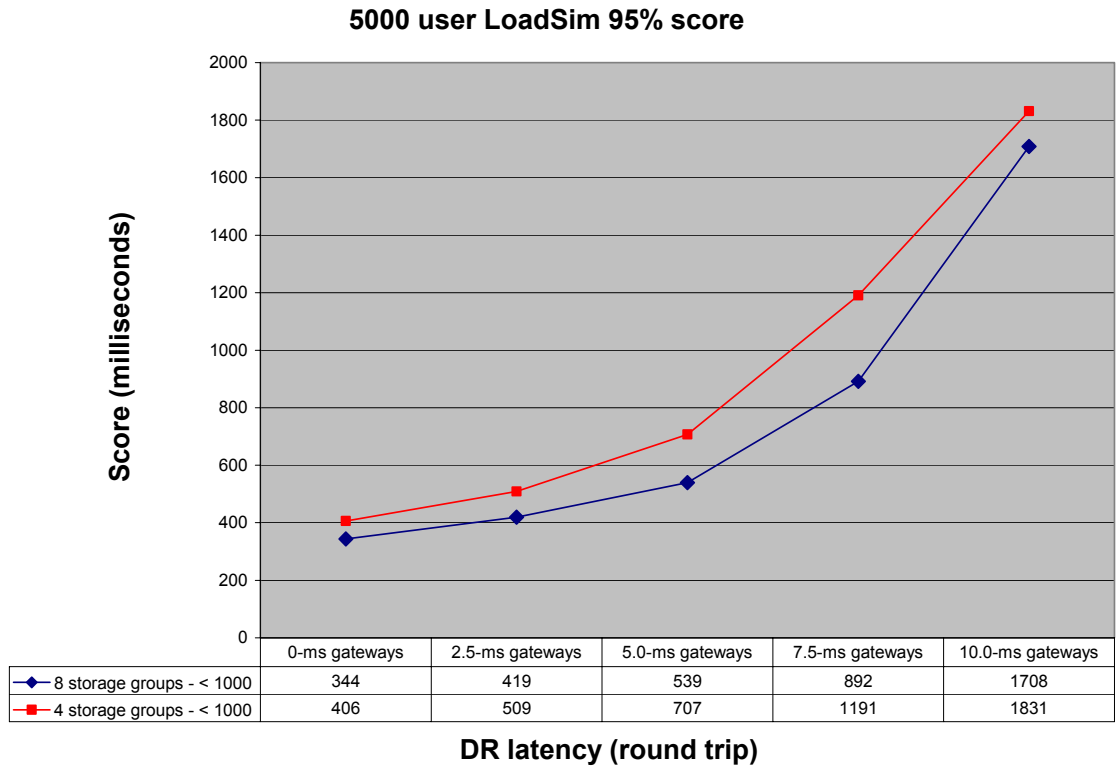
Average CPU usage was constant, as shown in Figure 6. A constant CPU workload is desirable from a LoadSim perspective. Overloading a LoadSim client server would not allow proper simulation of clients on that server. In other words, the server would not generate the proper message load on the Exchange virtual servers.

Figure 6 also shows that average client response time, while increasing, remained well below the acceptable LoadSim limits of 1000 ms. However, it was common to see maximum values spike above 1000 ms. This further shows that Exchange clients would see reasonable response times as latencies increased. The data shows that there were not any prolonged response delays for the clients.

Figure 6. LoadSim CPU usage and client request latency

5000 user LoadSim client

	0-ms gateways	2.5-ms gateways	5.0-ms gateways	7.5-ms gateways	10.0-ms gateways
8 SG CPU usage	31.75	31.875	30.375	32	31.375
4 SG CPU usage	32	32.625	31.875	32.375	29.5
8 SG latency - < 1000	91.75	101.5	115.5	167.375	276.2857143
4 SG latency - < 1000	107.375	123.625	153.75	215.25	369.5

DR latency (round trip)
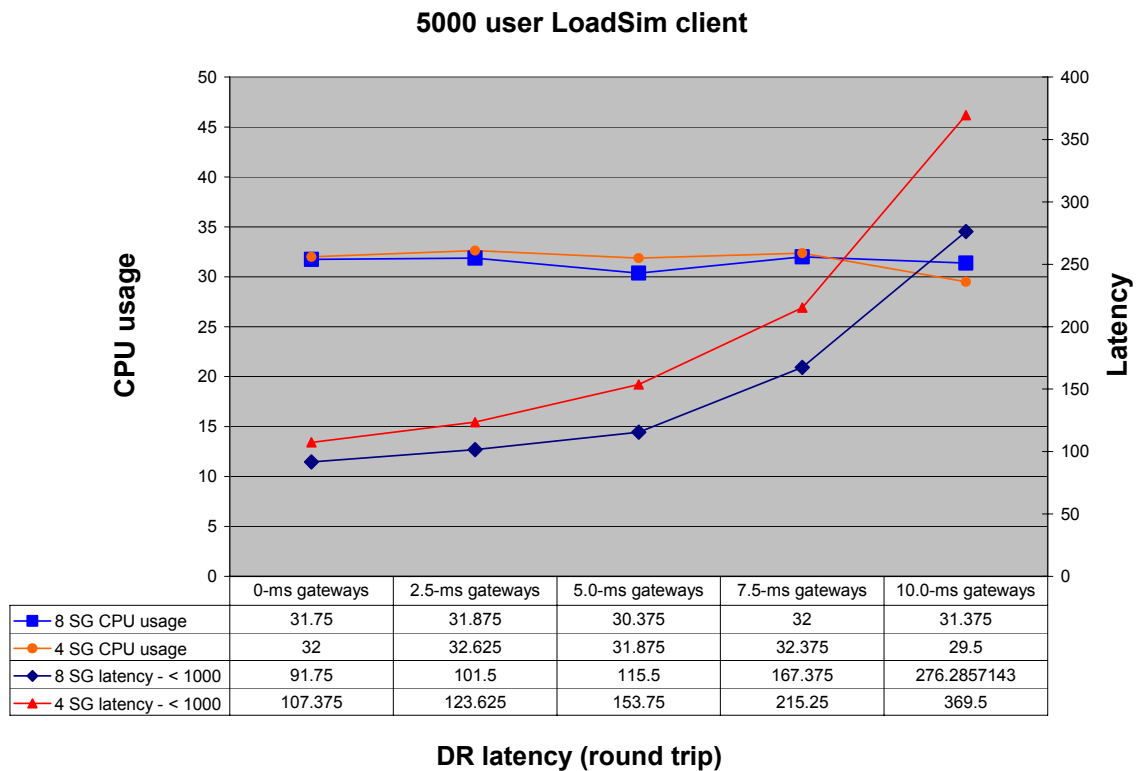
Phases 1 and 2 LoadSim Exchange Server performance data

Test results indicate the CPU usage of the Exchange virtual servers in this configuration were not greatly affected by longer inter-site latencies. Results showed the Exchange servers continued to process client requests in a timely manner.

As a reference, the CPU load created by the testing in Phases 1 and 2 testing did not push the Exchange servers. As shown in Figure 7, CPU usage remained constant at approximately 40%, even as DR latency increased. The goal of this testing was to determine the effect of increased inter-site latencies, not to push the limits of the CPUs.

Figure 7. Exchange server CPU usage

5000 user EVS CPU usage

	0-ms gateways	2.5-ms gateways	5.0-ms gateways	7.5-ms gateways	10.0-ms gateways
8 SG CPU usage	38	39.5	40	40	45.1
4 SG CPU usage	38	40.4	40.45	41	37.6

DR latency (round trip)
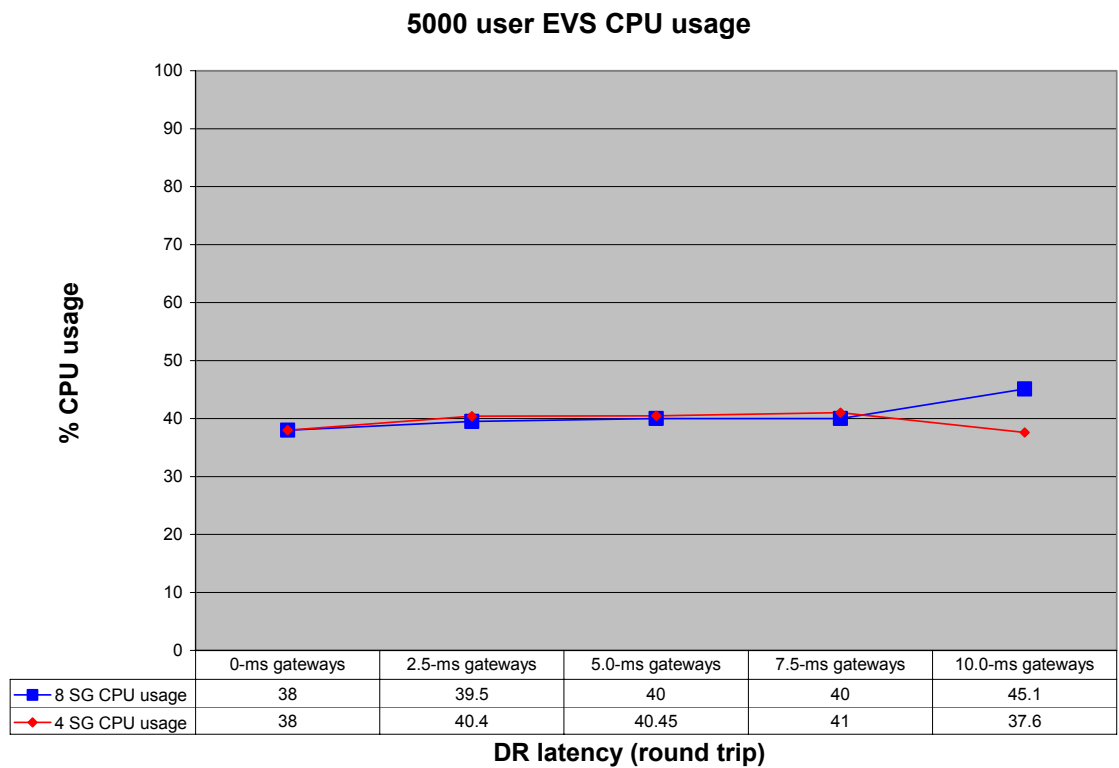
Even as DR latency increased, clients were not affected in their ability to submit or receive messages. Messaging performance remained constant, as shown in Figure 8. Results show values achieved above target values. Message performance is measured by the following LoadSim performance counters:

- **MSExchangeIS "Message recipients delivered per minute"**—Shows the rate at which recipients receive messages in the mailbox store. For a given user workload, a predicted message receive rate can be calculated. The actual value should be the target value or higher. The target value for 2,500 users is:

    (0.4375) * (number of users) = 0.4375 * 2,500 users per server = 1094

- **MSExchangeIS "Messages submitted per minute"**—Shows the rate at which messages are submitted by clients. For a given user workload, a predicted message submission rate can be calculated. The actual value should be the target value or higher. The target value for 2,500 users is:

    (number of users) * (0.1771) = 2,500 users per server * 0.1771 = 442

As can be seen in Figure 8, the actual values were higher than the target for both performance counters out to inter-site latencies of 7.5 ms. DR latencies longer than 7.5 ms affected the Phase 2 testing with four storage groups to a larger degree.

Figure 8. Exchange message delivery rates

5000 user EVS message

	0-ms gateways	2.5-ms gateways	5.0-ms gateways	7.5-ms gateways	10.0-ms gateways
8 SG msgs sub/min >= 442	538.5	536	530.5	514	477.5
8 SG msg recipients del/min >= 1093	1114	1142.5	1133	1128.5	974.5
4 SG msgs sub/min >= 442	480	487.5	456	468.5	353.5
4 SG msg recipients del/min >= 1093	1117.5	1145	1109.5	1120	669

DR latency (round trip)
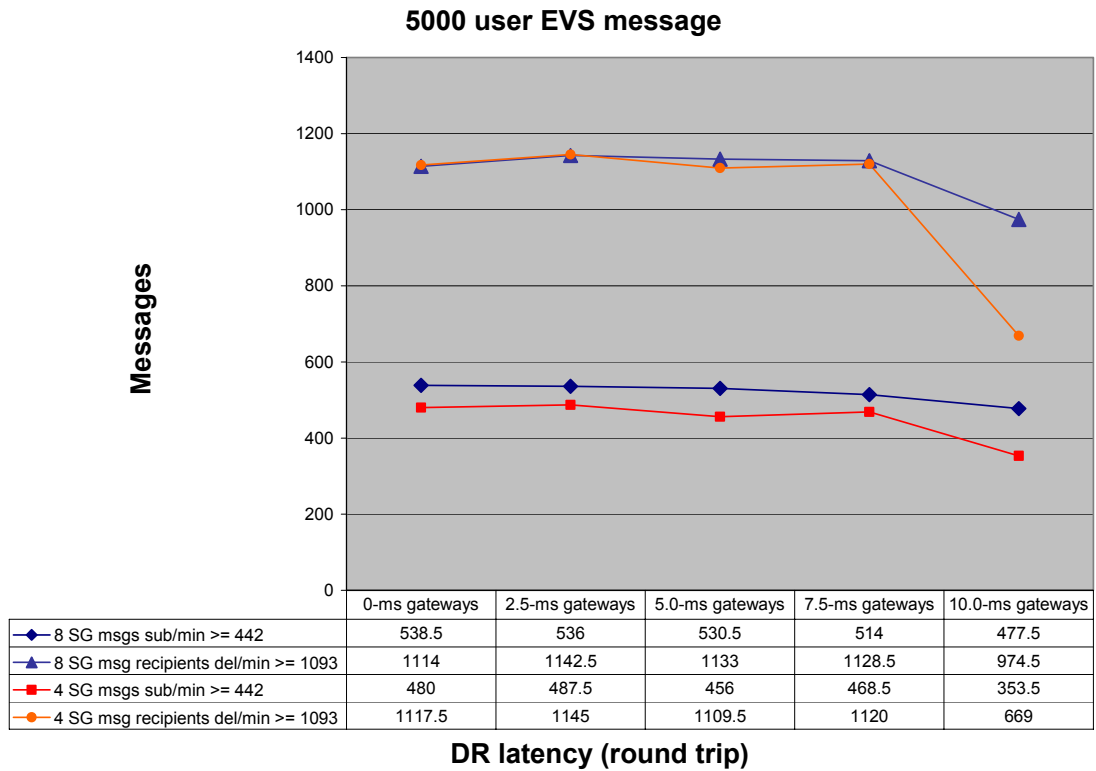
Message traffic between Exchange clients and servers occurs through the use of remote procedure calls (RPC). Typically, RPC requests occur asynchronously. Any delay by the Exchange server in processing these requests directly affects user experience and system responsiveness.

Figure 9 shows that the average RPC request latency increased as DR latency increased. RPC latency for Phase 1 testing with eight storage groups remained below 50 ms up to the 8-ms inter-site latency. Phase 2 testing with four storage groups crossed that threshold sooner at the 7-ms DR latency value.
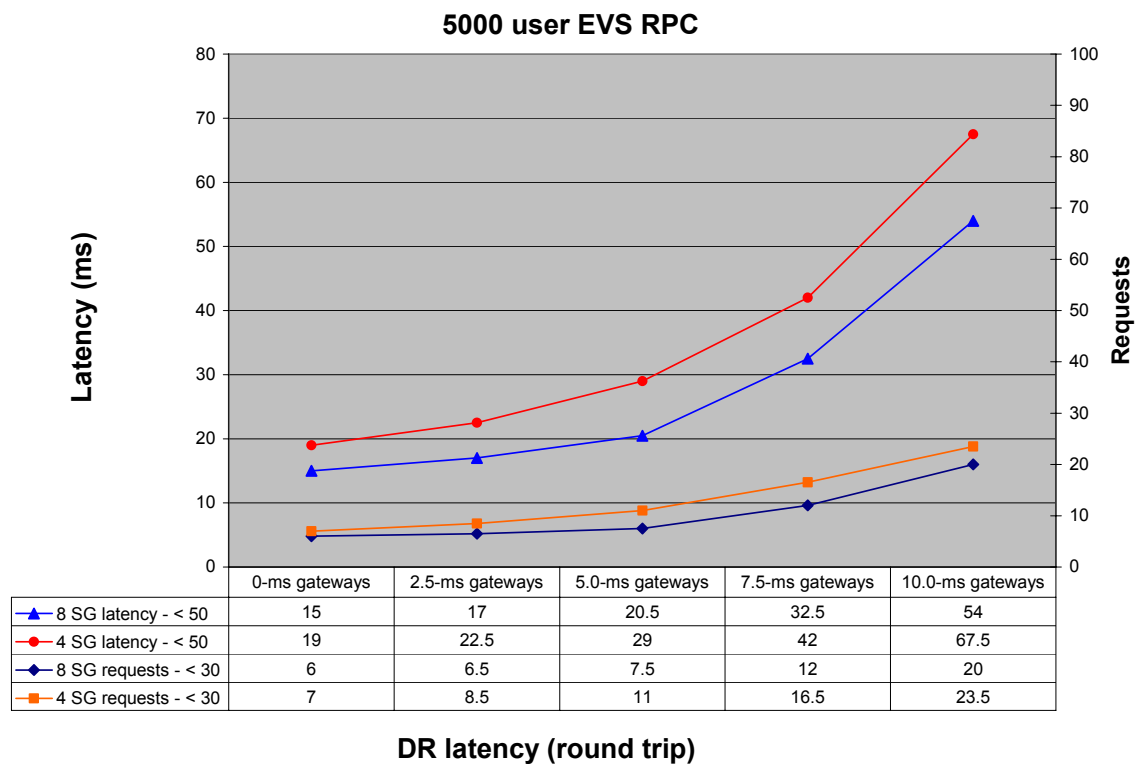
The number of RPC requests being serviced at any one time by the Exchange server remained fairly constant and remained below the 30 request limit as inter-site latencies increased.

The RPC request and latency values show that the Exchange virtual servers were processing client requests in a timely manner out to the Phase 1 and 2 latencies of 8 ms and 7 ms, respectively. The Exchange virtual servers remained responsive to users up to these DR latency values.

Figure 9. Exchange RPC

5000 user EVS RPC

	0-ms gateways	2.5-ms gateways	5.0-ms gateways	7.5-ms gateways	10.0-ms gateways
8 SG latency - < 50	15	17	20.5	32.5	54
4 SG latency - < 50	19	22.5	29	42	67.5
8 SG requests - < 30	6	6.5	7.5	12	20
4 SG requests - < 30	7	8.5	11	16.5	23.5

DR latency (round trip)

Exchange performance can degrade if one or more mail queues begin to build up.

Simple Mail Transport Protocol (SMTP) is the default transport for messages across servers and to the Internet. A buildup in the SMTP local queue indicates the Exchange server can no longer deliver the incoming mail in a timely manner. A buildup in the SMTP remote queue means that mail is not being sent to other servers.

Figure 10 shows that the SMTP local queue length increased as DR latency increased but remained below the 1000-ms limit for Phase 1 and 2 testing up to inter-site latencies of approximately 8.7 ms. The SMTP remote queue length showed a corresponding increase as inter-site latencies increased.

Figure 10. Exchange SMTP local and remote server queues

5000 user EVS SMTP local and remote queues

	0-ms gateways	2.5-ms gateways	5.0-ms gateways	7.5-ms gateways	10.0-ms gateways
8 SG local queue - < 1000	20	33	51.5	119.5	3159.5
4 SG local queue - < 1000	25.5	39	60	147	14702.5
8 SG remote queue - < 1000	4	6	10	22	65.5
4 SG remote queue - < 1000	4.5	9.5	18	40	121

DR latency (round trip)
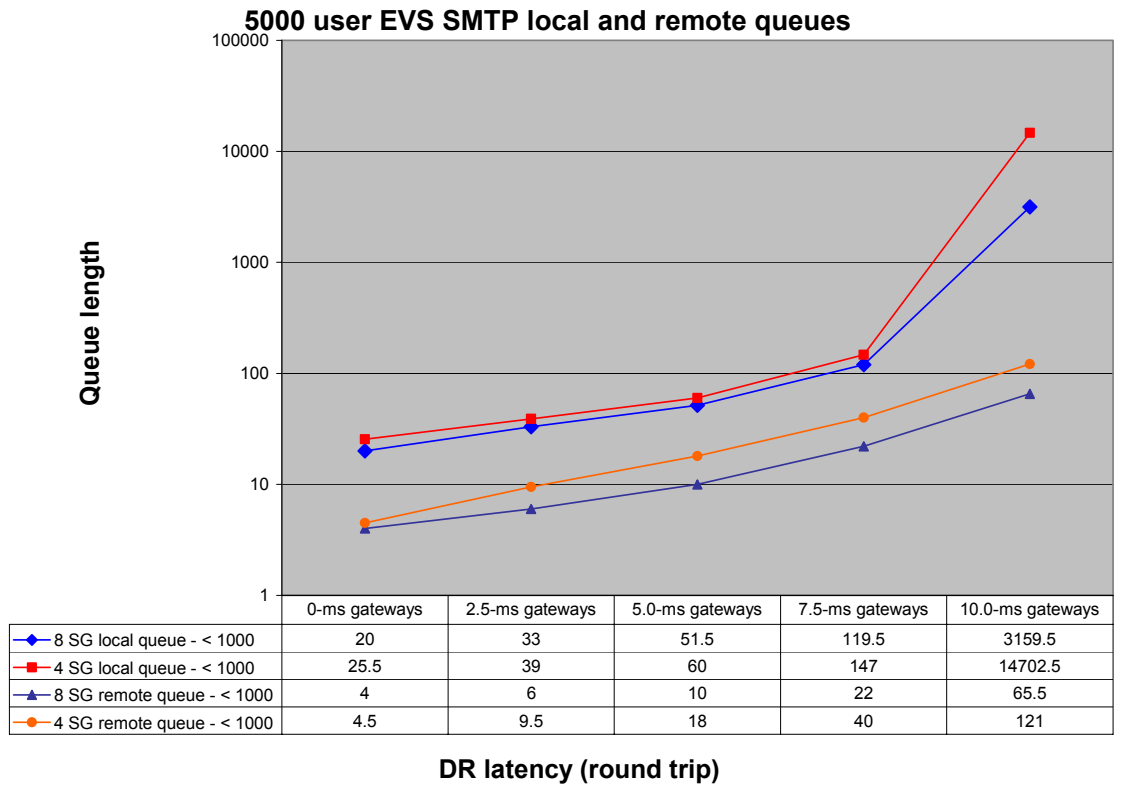
Figure 11 shows that the length of the Exchange Server Mailbox Stores send queue increased past the desired limit of 25 at approximately 5 ms of DR latency for Phase 1 and 2 testing. At 10 ms, the send queue length increased to approximately 190. The receive queue remained at zero at all times.

Figure 11. Exchange server mailbox send queue

5000 user EVS mailbox send queue

	0-ms gateways	2.5-ms gateways	5.0-ms gateways	7.5-ms gateways	10.0-ms gateways
8 SG send queue - < 25	11.5	18.5	29	63.5	192.5
4 SG send queue - < 25	14	24	38	73.5	189.5

Queue length

DR latency (round trip)
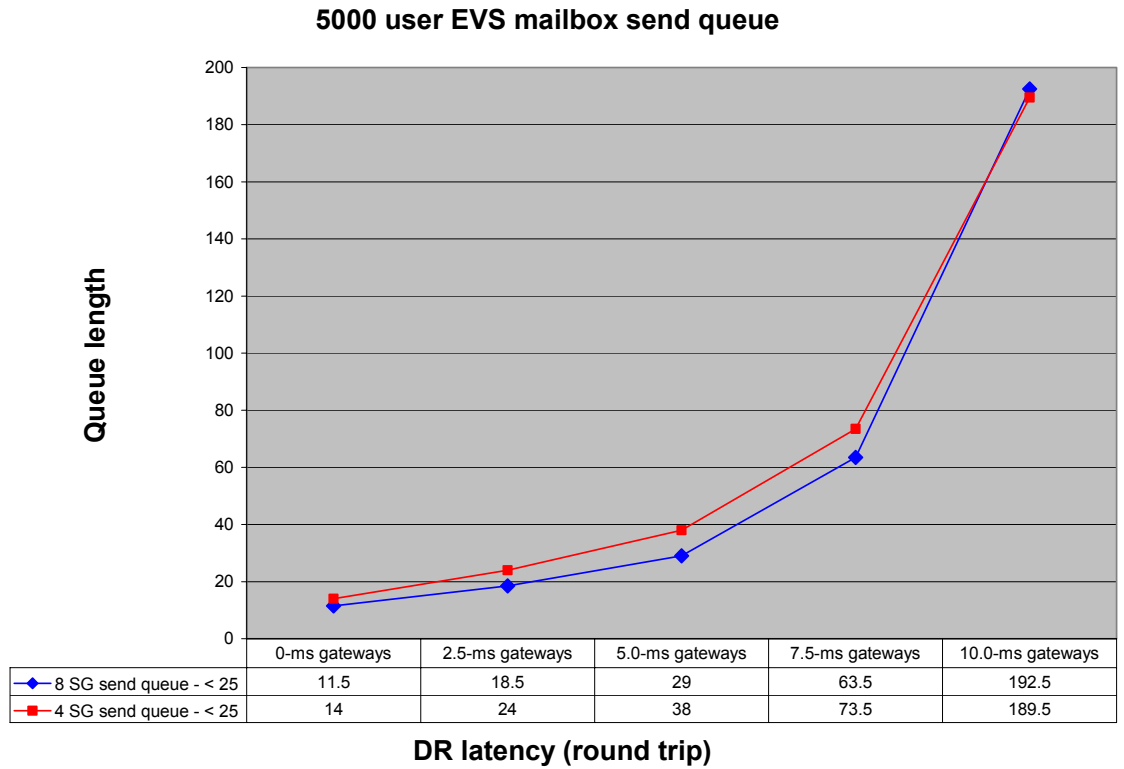
Phases 1 and 2 LoadSim disk performance

Exchange disk performance indicates that longer inter-site latency increases write latencies. Exchange is sensitive to increased latencies.

Log disk write latencies are shown in Figure 12. The write latency value crossed the 20-ms threshold at 8 ms and 7 ms, for Phases 1 and 2, respectively. L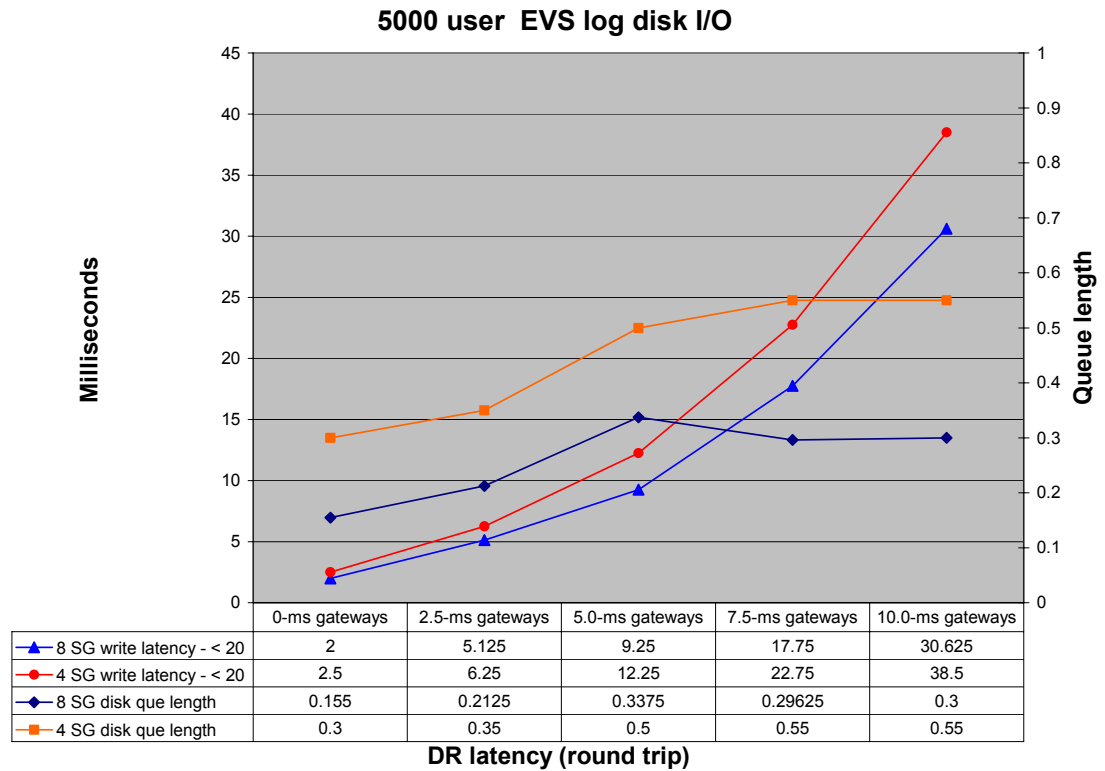og disk queue lengths remained low. Phase 1 testing shows that log write latency maximums remained below the 40-ms limit out to an inter-site latency of 8 ms, as shown in Figure 13.

Figure 12. Log disk I/O latencies

5000 user EVS log disk I/O

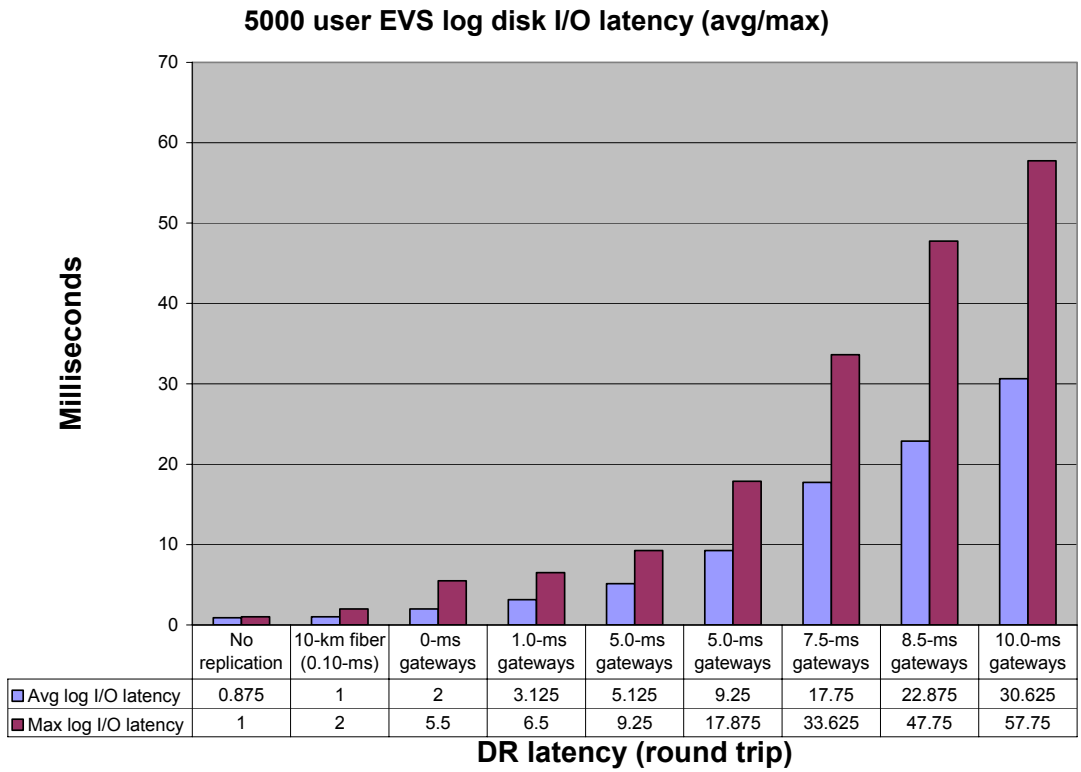|	0-ms gateways	2.5-ms gateways	5.0-ms gateways	7.5-ms gateways	10.0-ms gateways
8 SG write latency - < 20	2	5.125	9.25	17.75	30.625
4 SG write latency - < 20	2.5	6.25	12.25	22.75	38.5
8 SG disk que length	0.155	0.2125	0.3375	0.29625	0.3
4 SG disk que length	0.3	0.35	0.5	0.55	0.55

DR latency (round trip)

Note the peak in the 8-SG disk queue length line in Figure 12 at the 5.0-ms point of the graph. This peak occurred because the Exchange transaction log buffers were increased after this test run. Increasing the transaction buffer to 9000 reduced the transaction log disk I/Os, as seen in this graph.

Figure 13 shows the average and maximum log disk latency values, for reference. Typically, maximum values represent the largest spike of a given performance value during data collection.

Figure 13. Log I/O latency average and maximum for eight storage groups

5000 user EVS log disk I/O latency (avg/max)

	No replication	10-km fiber (0.10-ms)	0-ms gateways	1.0-ms gateways	5.0-ms gateways	5.0-ms gateways	7.5-ms gateways	8.5-ms gateways	10.0-ms gateways
Avg log I/O latency	0.875	1	2	3.125	5.125	9.25	17.75	22.875	30.625
Max log I/O latency	1	2	5.5	6.5	9.25	17.875	33.625	47.75	57.75

DR latency (round trip)

Database disk latencies also increased with longer inter-site latencies, as shown in Figure 14. The average database disk write latency was 130 ms for both Phase 1 and 2 at the 8-ms and 7-ms inter-site latencies. Database disk queues show a more dramatic increase for the Phase 2 testing.

Results show that the Phase 1 testing tolerated longer inter-site latencies better than Phase 2 testing. Phase 2 had more users per storage group, which doubled the load, causing the disks to become a bottleneck.

Figure 14. Database disk I/O latencies

5000 user EVS DB disk I/O

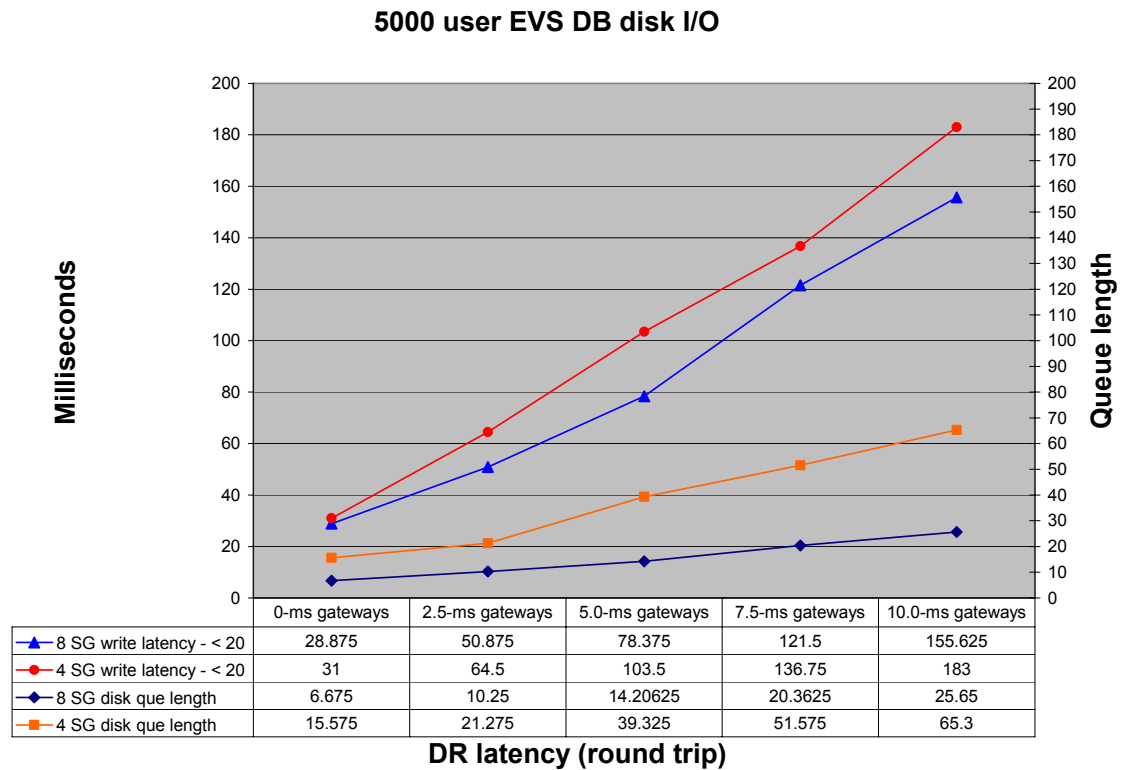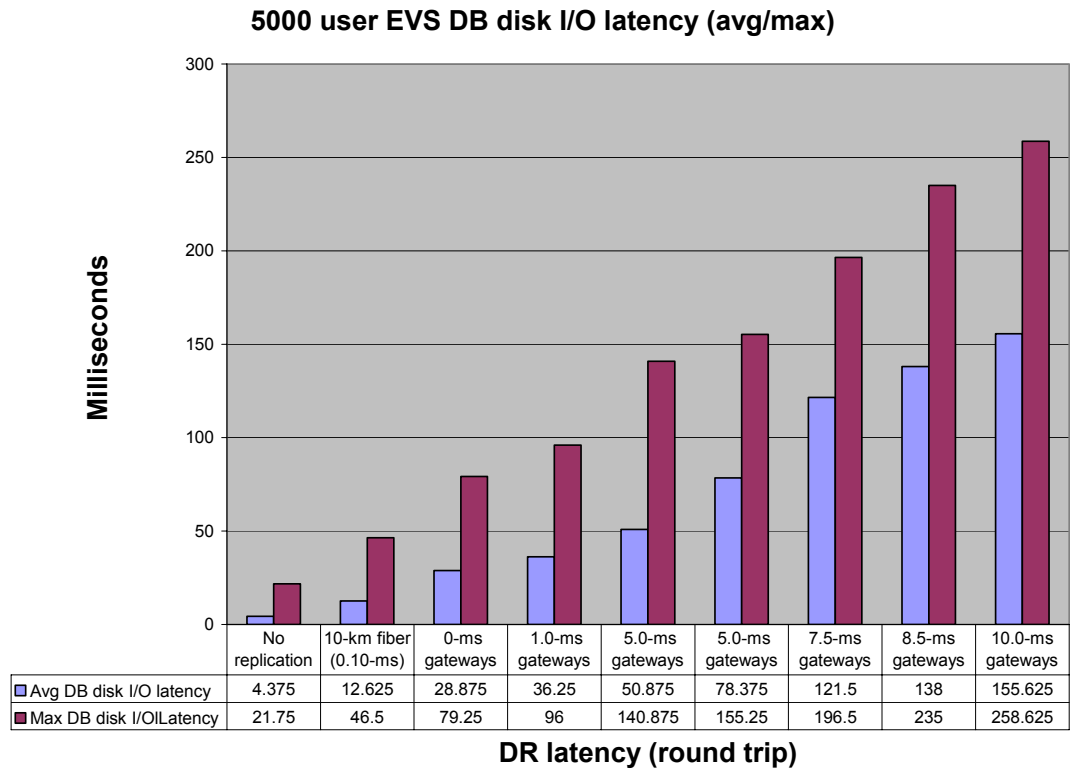|	0-ms gateways	2.5-ms gateways	5.0-ms gateways	7.5-ms gateways	10.0-ms gateways
8 SG write latency - < 20	28.875	50.875	78.375	121.5	155.625
4 SG write latency - < 20	31	64.5	103.5	136.75	183
8 SG disk que length	6.675	10.25	14.20625	20.3625	25.65
4 SG disk que length	15.575	21.275	39.325	51.575	65.3

DR latency (round trip)

Figure 15 shows the average and maximum database disk latency values for Phase 1 testing. The maximum database disk latency values are greater than 150 ms at a 10-ms inter-site latency.

Figure 15. Database disk I/O latency average and maximum for four storage groups

5000 user EVS DB disk I/O latency (avg/max)

	No replication	10-km fiber (0.10-ms)	0-ms gateways	1.0-ms gateways	5.0-ms gateways	5.0-ms gateways	7.5-ms gateways	8.5-ms gateways	10.0-ms gateways
Avg DB disk I/O latency	4.375	12.625	28.875	36.25	50.875	78.375	121.5	138	155.625
Max DB disk I/OlLatency	21.75	46.5	79.25	96	140.875	155.25	196.5	235	258.625

DR latency (round trip)

Phase 1 Jetstress performance testing

The Jetstress testing further validated the LoadSim testing results: increased inter-site latency affects disk I/O performance.

The Jetstress testing was divided into two parts. The first consisted of Jetstress data replication latency testing over latencies of 0.10 ms to 5.0 ms to determine the effect of increased inter-site latency. The second part consisted of Jetstress I/O throughput testing, where the goal was to increase the load from 1,000 IOPS to 12,000 IOPS over DR latencies of 0.10 ms and 1.0 ms.

Jetstress testing
Jetstress configuration

For Jetstress testing, the JetstressUI tool was used. The configuration setup is shown in Figure 16 and Figure 17. The configuration data shown is for one server. However, both the Exchange virtual servers (EVS) were tested. The workload selected was moderate, 0.5 IOPS per 100-MB mailbox. A total of 2,500 mailboxes per EVS were tested. Performance runs lasted two hours. Ten Jetstress threads were used to generate the workload.
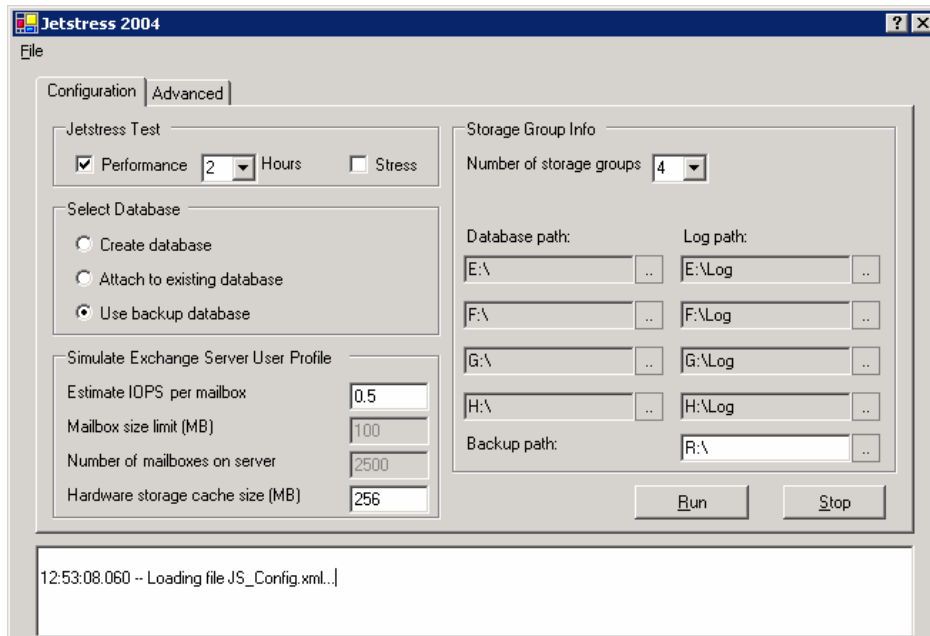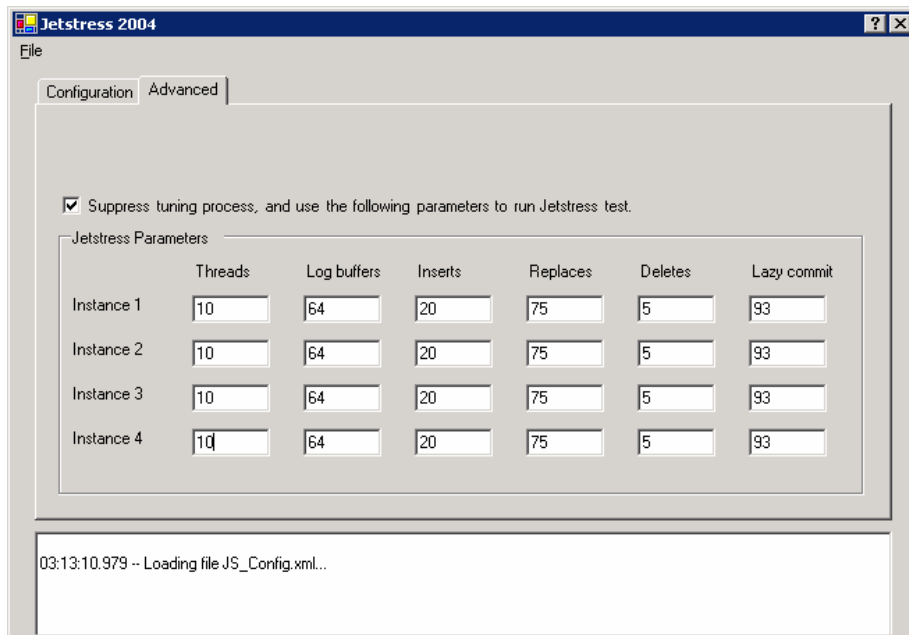
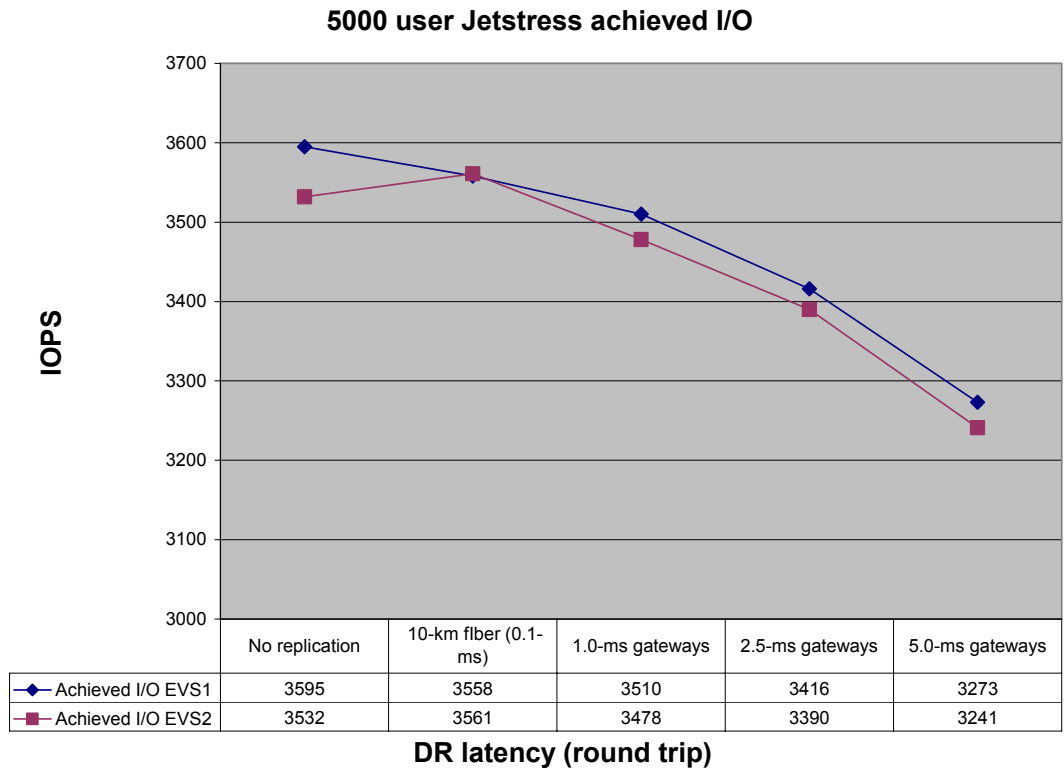**Figure 16.** Jetstress configuration

Figure 17. Jetstress tuning parameters
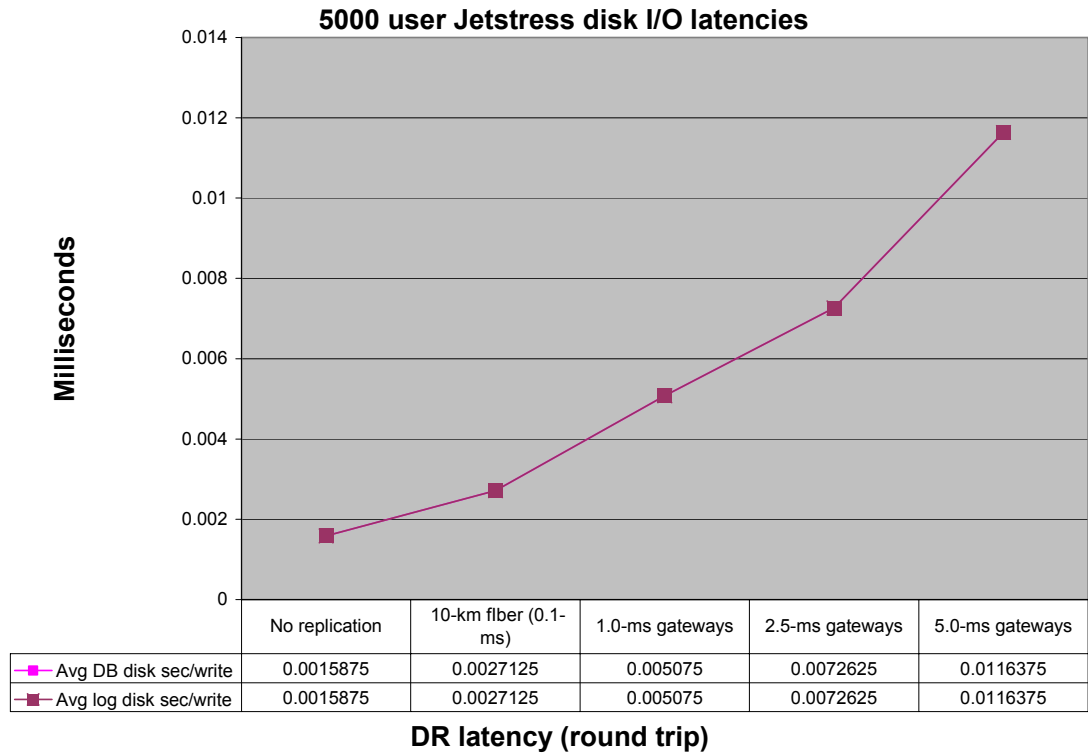
Jetstress performance data

Jetstress testing showed similar results to the LoadSim testing. As the inter-site latency increased, I/O latencies increased. Figure 18 shows that the disk I/O throughput decreases by approximately 8% with a 5.0-ms inter-site latency, compared with no data replication.

Figure 18. Jetstress achieved I/O

5000 user Jetstress achieved I/O

	No replication	10-km fiber (0.1-ms)	1.0-ms gateways	2.5-ms gateways	5.0-ms gateways
Achieved I/O EVS1	3595	3558	3510	3416	3273
Achieved I/O EVS2	3532	3561	3478	3390	3241

DR latency (round trip)

IOPS

Disk write latencies increased from 2 ms to just less than 12 ms, as shown in Figure 19.

Figure 19. Jetstress disk I/O latencies

5000 user Jetstress disk I/O latencies

	No replication	10-km fiber (0.1-ms)	1.0-ms gateways	2.5-ms gateways	5.0-ms gateways
Avg DB disk sec/write	0.0015875	0.0027125	0.005075	0.0072625	0.0116375
Avg log disk sec/write	0.0015875	0.0027125	0.005075	0.0072625	0.0116375

DR latency (round trip)

Phase 1 Jetstress throughput testing

The goal was to run Jetstress and apply increasing loads of 1,000, 4,000, 8,000, and 12,000 IOPS. Throughput testing used the following test parameters:

- 1.0 IOPS per mailbox
- A change in the number of threads to increase the I/O load. For this testing, the number of threads tested was 1, 5, 10, 20, and 40.
- Performance test runs at 10-km (0.1-ms) and 1.0-ms latencies.

Results show that a maximum I/O throughput of 4,000 was achieved for both latencies in this configuration.

Figure 20 and Figure 21 show disk latencies increasing as the number of threads increases.

Figure 20. Jetstress disk latencies—10 km (0.10-ms inter-site latency)
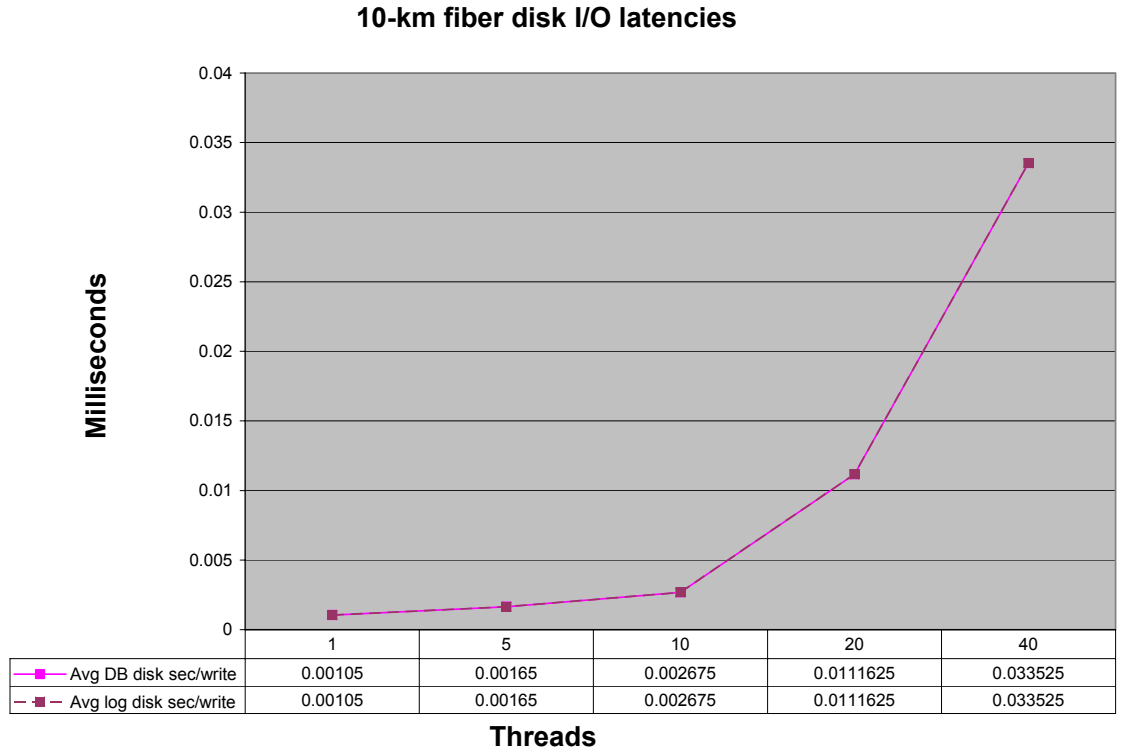

10-km fiber disk I/O latencies

	1	5	10	20	40
Avg DB disk sec/write	0.00105	0.00165	0.002675	0.0111625	0.033525
Avg log disk sec/write	0.00105	0.00165	0.002675	0.0111625	0.033525

Threads

Figure 21. Jetstress disk latencies—1.0-ms inter-site latency

1.0-ms DR latency disk I/O latency

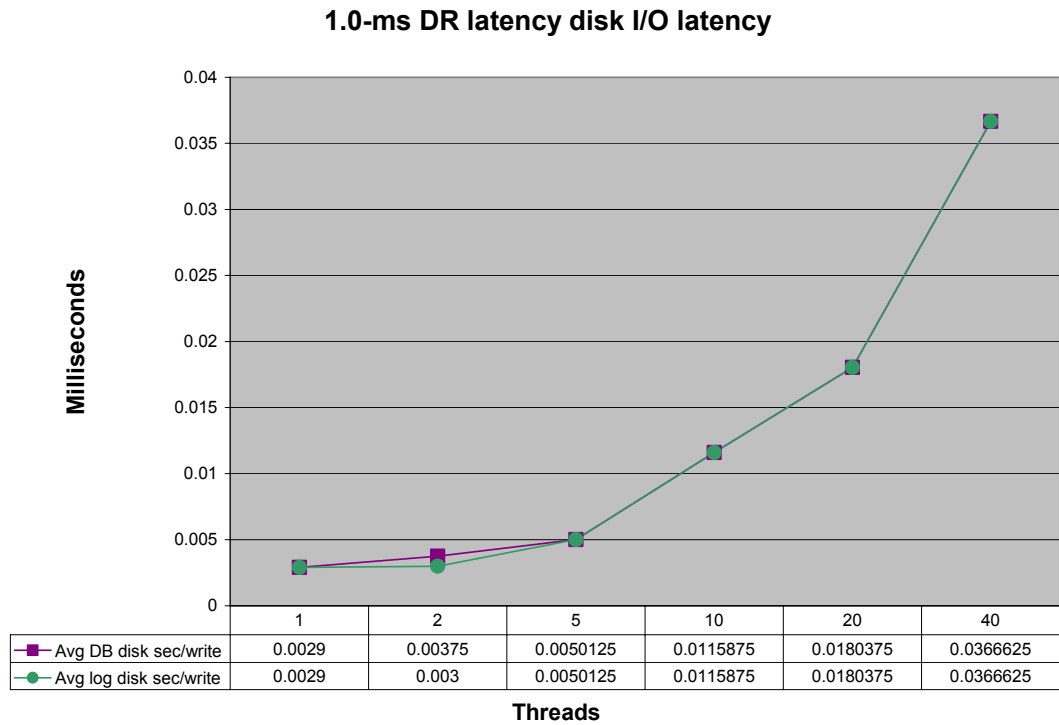|	1	2	5	10	20	40
Avg DB disk sec/write	0.0029	0.00375	0.0050125	0.0115875	0.0180375	0.0366625
Avg log disk sec/write	0.0029	0.003	0.0050125	0.0115875	0.0180375	0.0366625

Threads

Figure 22 and Figure 23 show that achieved I/O throughput did not increase for the 40-thread test run. The test run data for 10 km (0.1 ms) was similar to that for the 1.0-ms DR latency.

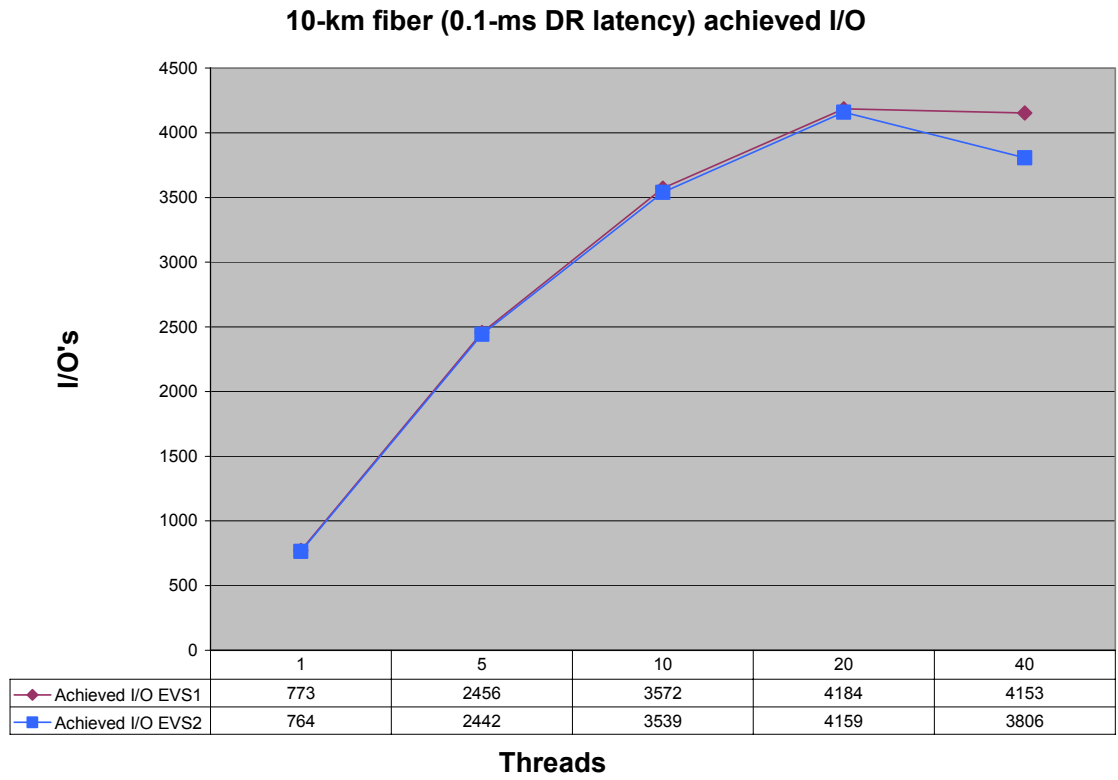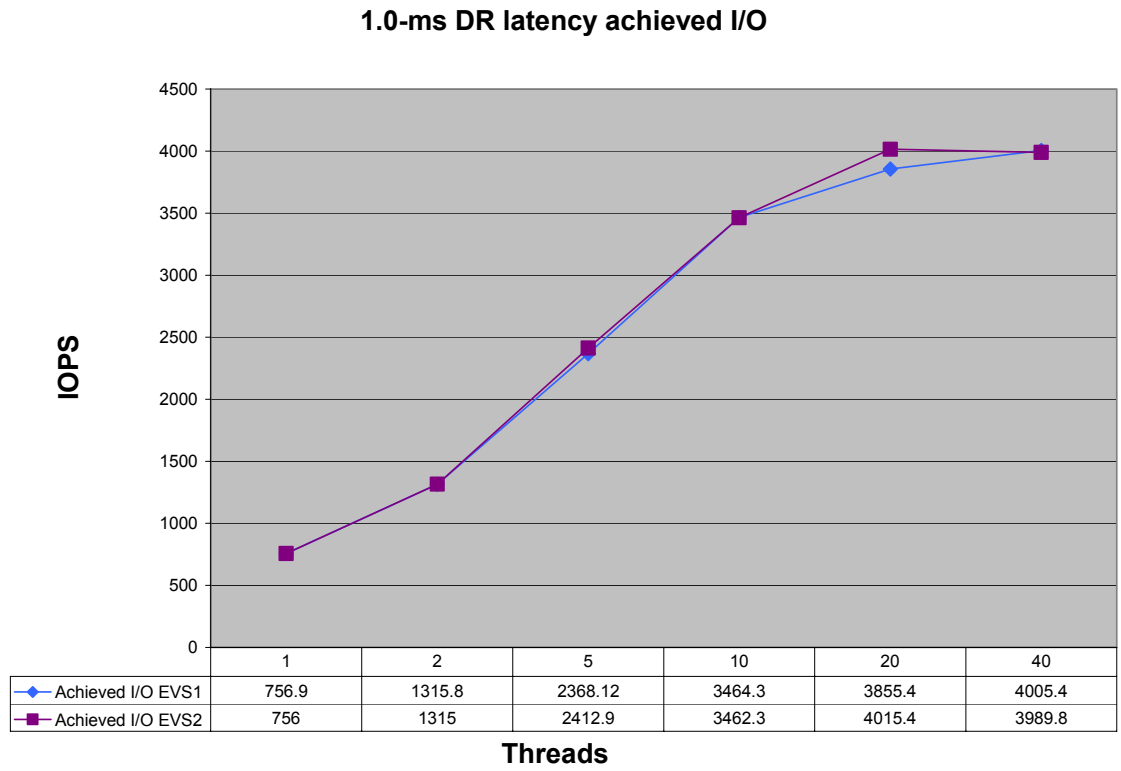**Figure 22.** Jetstress achieved I/O throughput—10 km

10-km fiber (0.1-ms DR latency) achieved I/O

	1	5	10	20	40
Achieved I/O EVS1	773	2456	3572	4184	4153
Achieved I/O EVS2	764	2442	3539	4159	3806

Threads

I/O's

Figure 23. Jetstress achieved I/O throughput—1.0-ms latency

1.0-ms DR latency achieved I/O

	1	2	5	10	20	40
Achieved I/O EVS1	756.9	1315.8	2368.12	3464.3	3855.4	4005.4
Achieved I/O EVS2	756	1315	2412.9	3462.3	4015.4	3989.8

IOPS

Threads

Phase 3 increased user load performance results

This testing consisted of executing three LoadSim test runs with a 1.0-ms inter-site latency with 5,000, 7,500, and 10,000 users against eight storage groups (same as Phase 1 testing). The goal was to see how Exchange reacted to increased workloads. The tests showed that as the user load increased, the two Exchange servers were finally overloaded and could not process incoming requests. The Exchange servers, not the storage, were the bottleneck.

Figure 24 and Figure 25 show that the log and database disks were not the bottleneck even though latencies and queue lengths are increasing. This is expected with increasing workload. However, log disk I/O latency did remain below the 20-ms threshold with a 10,000-user work load.
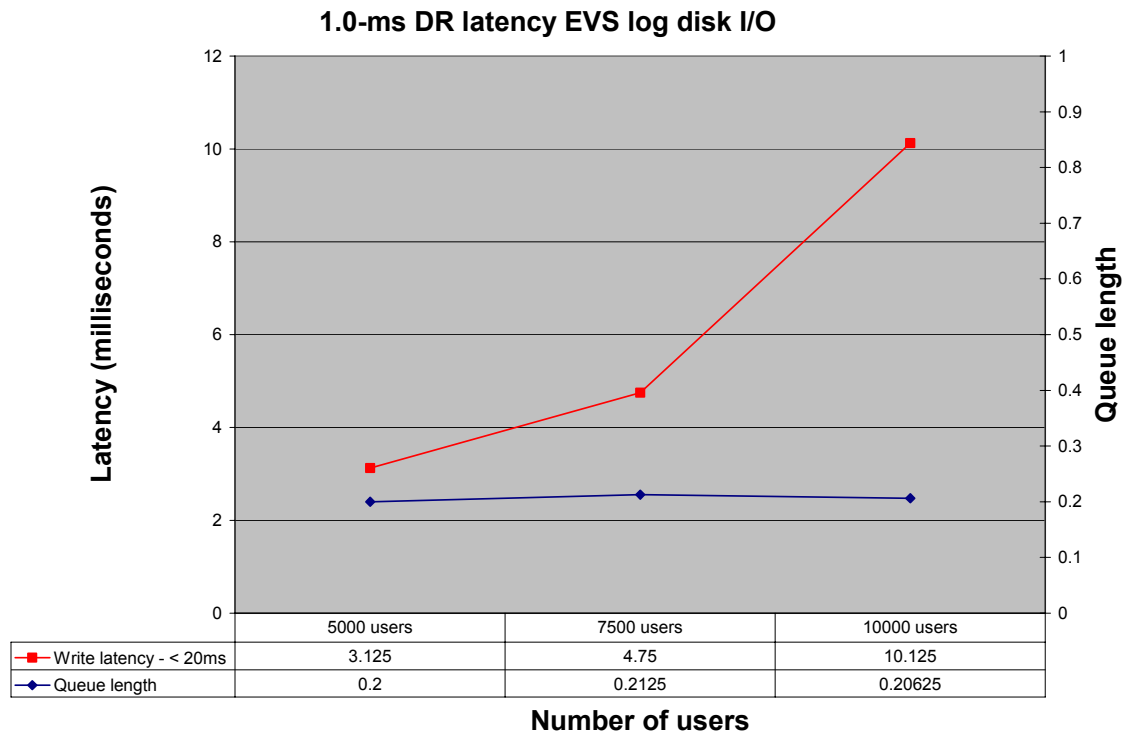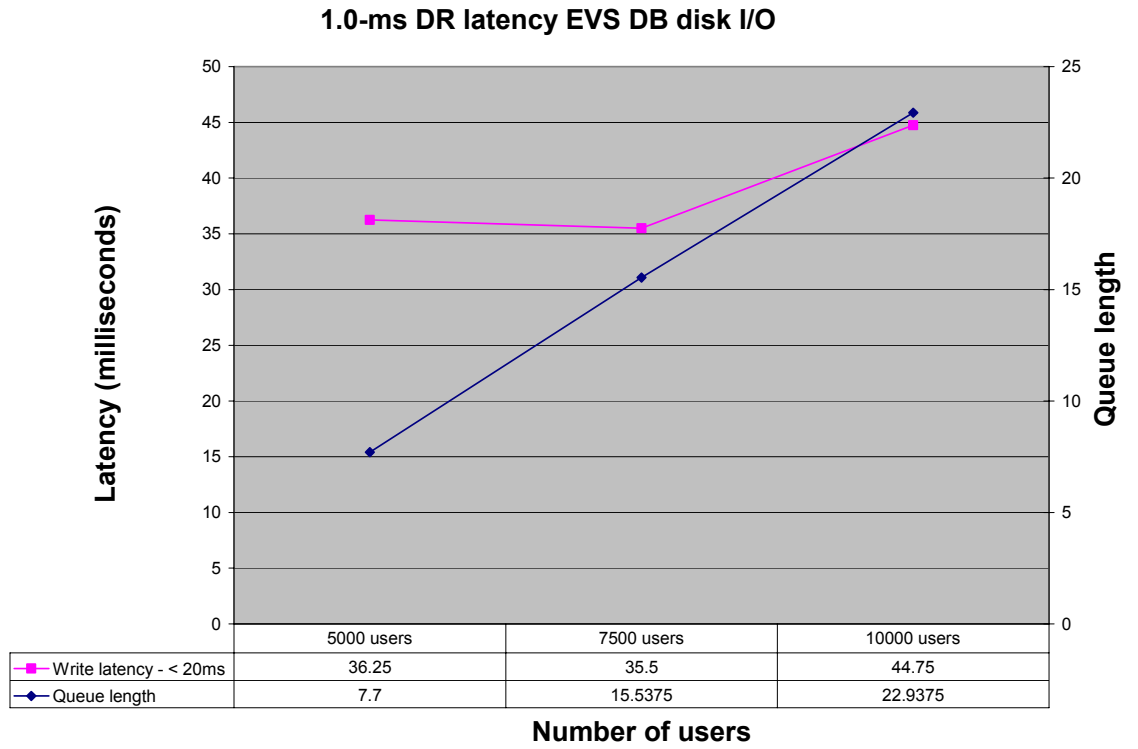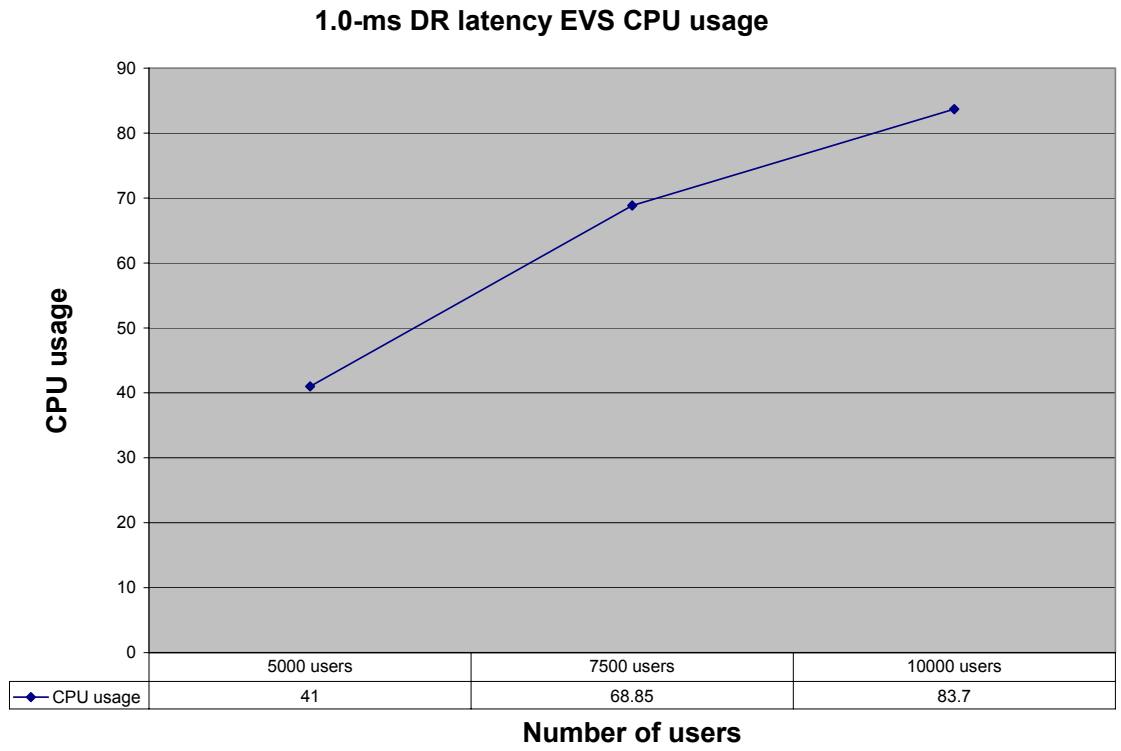
Figure 24. Log Disk I/O Latency

1.0-ms DR latency EVS log disk I/O

	5000 users	7500 users	10000 users
Write latency - < 20ms	3.125	4.75	10.125
Queue length	0.2	0.2125	0.20625

Number of users

Figure 25. Database Log Disk I/O

1.0-ms DR latency EVS DB disk I/O

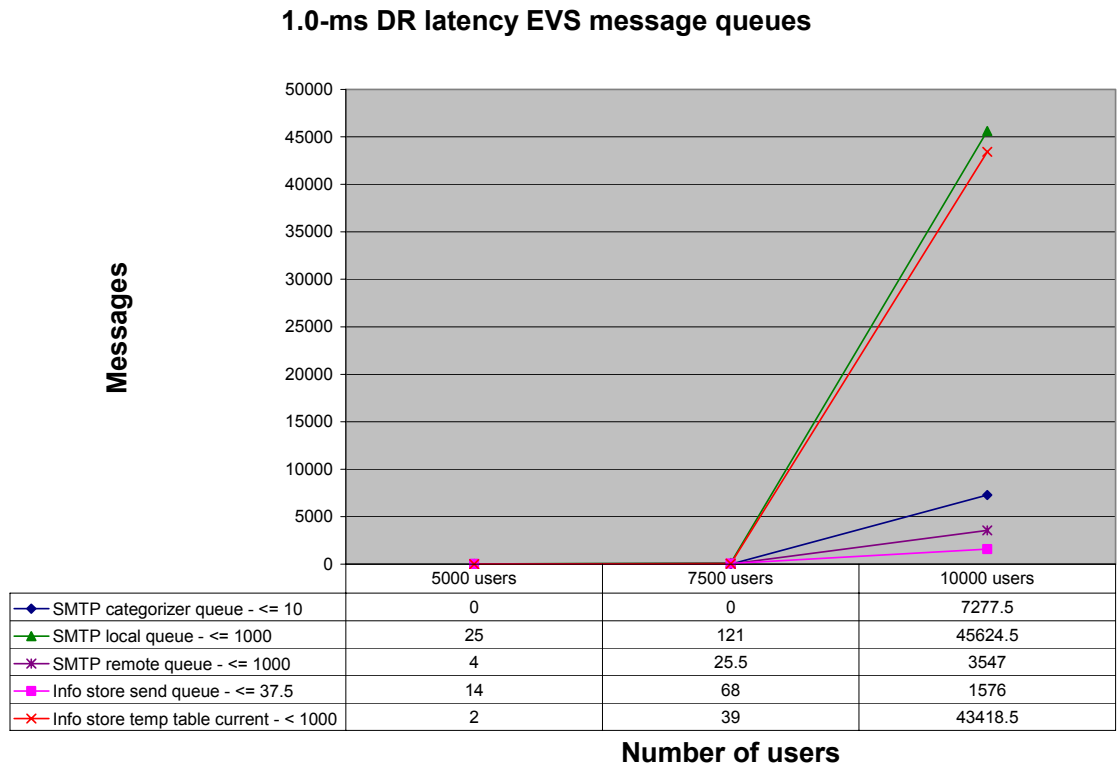|	5000 users	7500 users	10000 users
Write latency - < 20ms	36.25	35.5	44.75
Queue length	7.7	15.5375	22.9375

Increased user load also pushed the CPU usage to above 80%, as shown in Figure 26. CPU usage above 80% prevents optimal Exchange server performance.
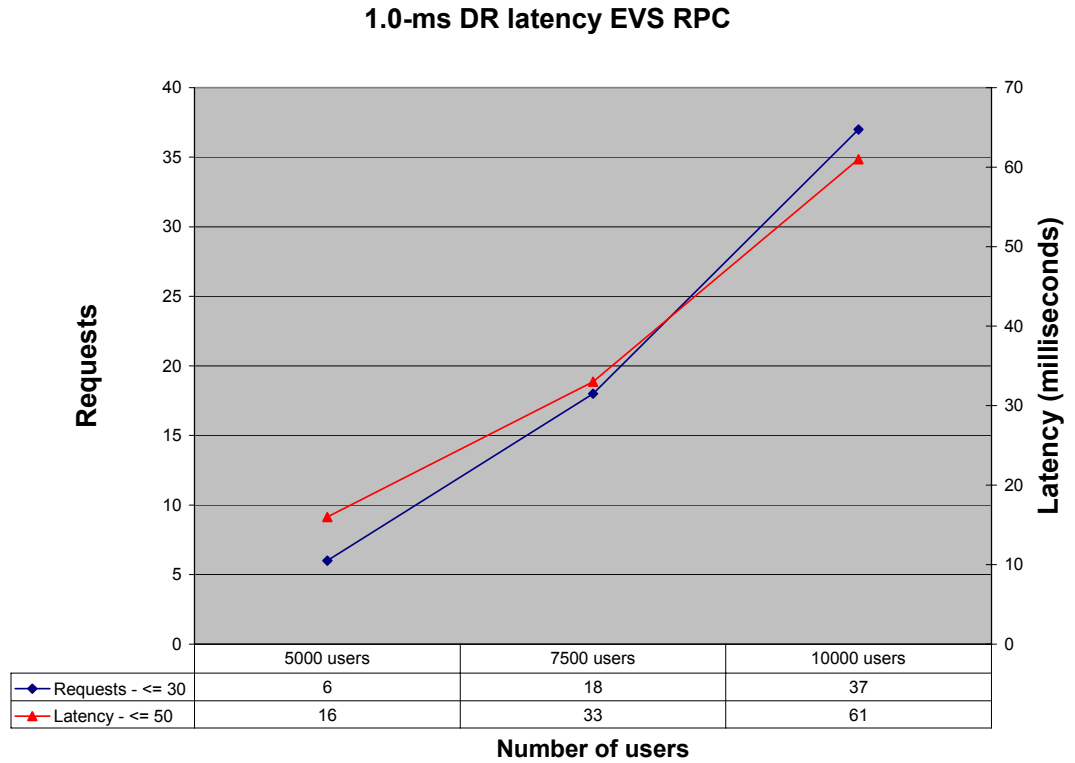
Figure 26. EVS CPU usage

1.0-ms DR latency EVS CPU usage

	5000 users	7500 users	10000 users
CPU usage	41	68.85	83.7

Number of users

The internal Exchange server queues became severely backed up with 10,000 users. Figure 27 shows queue lengths increasing as the Exchange servers were unable to keep up with demand.
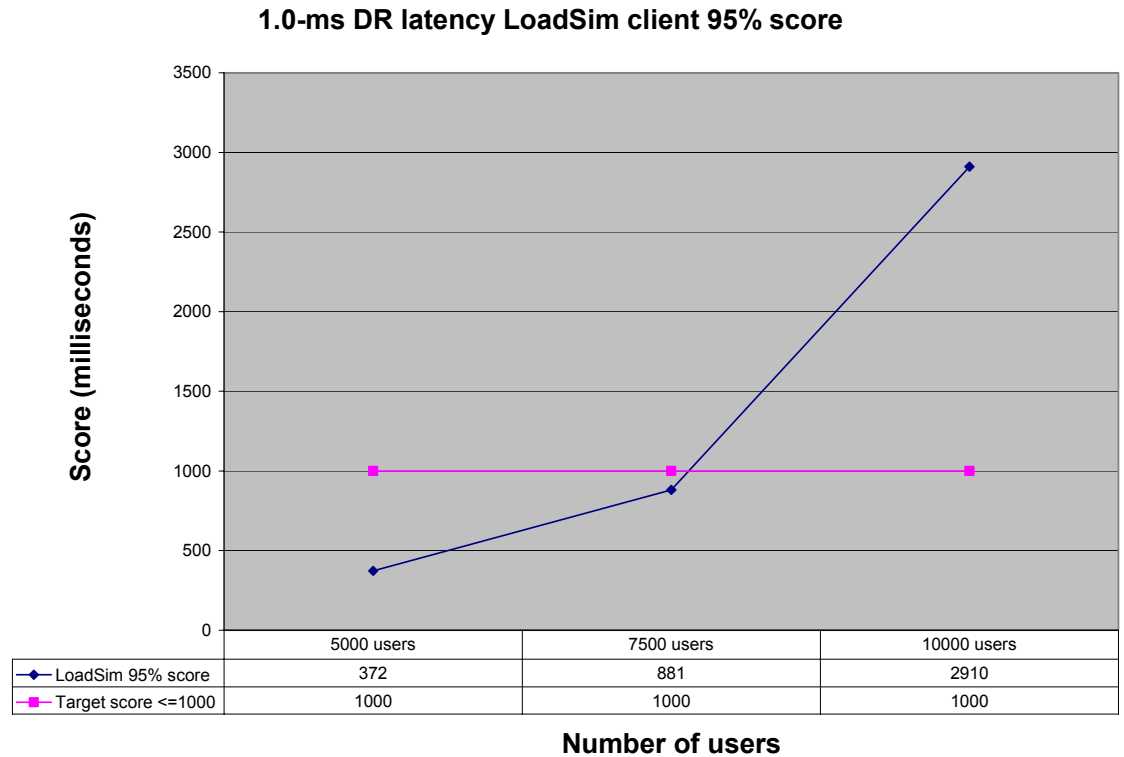
Figure 27. EVS message queues

1.0-ms DR latency EVS message queues

	5000 users	7500 users	10000 users
SMTP categorizer queue - <= 10	0	0	7277.5
SMTP local queue - <= 1000	25	121	45624.5
SMTP remote queue - <= 1000	4	25.5	3547
Info store send queue - <= 37.5	14	68	1576
Info store temp table current - < 1000	2	39	43418.5

Number of users

As the servers became overloaded, the ability of the servers to process requests was compromised. Figure 28 shows a dramatic increase in the number of active requests as the number of users increased. Correspondingly, the average time to process each request increased also. At 10,000 users, the RPC latency was above the 50-ms threshold.

Figure 28. EVS RPC

1.0-ms DR latency EVS RPC

	5000 users	7500 users	10000 users
Requests - <= 30	6	18	37
Latency - <= 50	16	33	61

Number of users

The "clients" in the 10,000-user test would have experienced a significant degradation in performance. The 95% score was triple the 1000-ms limit, as shown in Figure 29. Users would have experienced long delays and lost connections to the Exchange server.

Figure 29. LoadSim client 95th percentile score

1.0-ms DR latency LoadSim client 95% score

	5000 users	7500 users	10000 users
LoadSim 95% score	372	881	2910
Target score <=1000	1000	1000	1000

Number of users

Failover testing

Another objective of this testing was to demonstrate acceptable application failover with reduced failover times in the face of disasters. This is achieved by combining an HP StorageWorks Continuous Access data replication storage subsystem and a standby remote Exchange server cluster. Continuous Access data replication provides disaster tolerance by ensuring that the Exchange Information Stores are replicated at the remote site and provide real-time array-based synchronous replication. The remote Exchange server cluster is kept in a warm standby mode, allowing for immediate availability. What makes this design unique is that the local and remote Exchange server clusters are duplicates of each other. This configuration does not use a stretched cluster or an HP StorageWorks Cluster Extension EVA (CLX) solution.

Three failover simulations were tested in this configuration:

- Loss of storage (powering off the local EVA storage cabinet)
- Loss of site (powering off the local Exchange cluster and the primary domain controller servers)
- Loss of the ISL (see "Continuous Access write history log merge testing" on page 47 for details)

Cluster resources

The remote Exchange server cluster must be configured to match the local Exchange server cluster exactly. Figure 30 shows the cluster resources of the tested configuration. Of importance to this solution are the network name, IP address, and physical disk cluster resources. For ease of management, these resources have the same names at the local and remote sites. However, the network name and IP address resources are identical. The network name is what an Outlook client uses to connect to the correct mail server. The IP address is used as the actual network address. This solution requires that both local and remote clusters have the same network name and IP address, so that failover to the remote site is successful and clients are still able to connect to their mail server.

Figure 30. Cluster resources
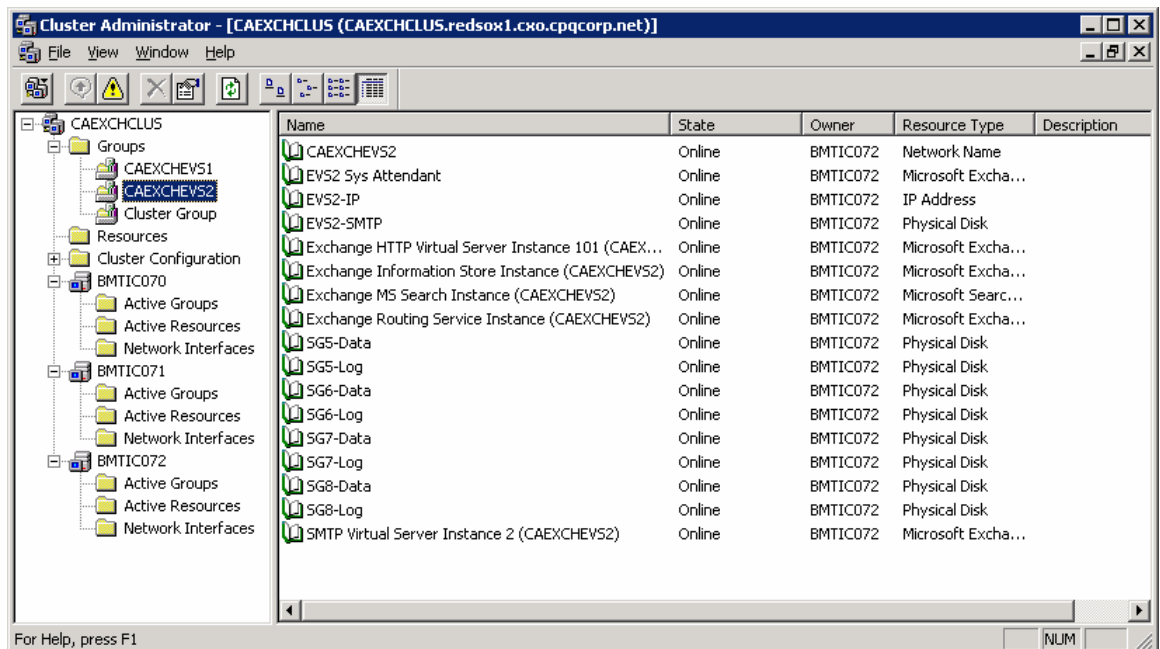
Remote Exchange server cluster warm standby mode

The remote Exchange server cluster waits in a warm standby mode until needed. Figure 31 shows the remote Exchange server cluster in standby mode. As you can see, the Exchange virtual server resource groups, CAEXCHEVS1 and CAEXCHEVS2, are offline. As a result, all the associated resources are also offline. Waiting in warm standby ensures that the remote cluster:

• Does not attempt to mount the Information Stores before they are available

• Remains unavailable, as the Exchange server, until needed

The cluster resource group can remain online. The cluster resource group contains the cluster name, cluster IP address, quorum, and DTC resources.

Figure 31. Cluster in warm standby mode
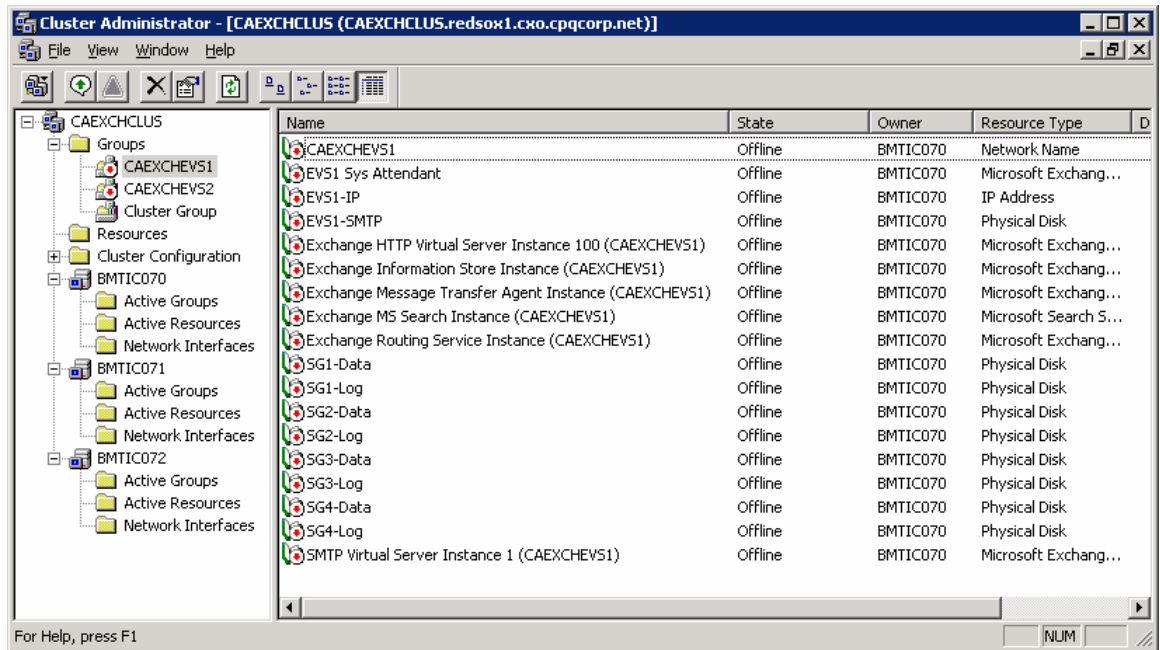
Failover testing results

During an EVA or site loss simulation, failover times of less than 15 minutes were achieved when the remote Exchange server cluster was in a warm standby mode. Broken down, the failover process was as follows.

Failover procedure		
Step	Action	Time to complete
1	Fail over Exchange Information Stores using Continuous Access user interface (CAUI).	1–3 minutes
2	Wait for Windows operating system to realize the storage disks have been presented.	1–2 minutes
3	Bring the Exchange virtual servers online. Most of this time was spent while the Exchange Information Store mounted all the Exchange Information Stores and performed database recovery.	3–5 minutes

The achieved failover time, in this case, would easily meet the 15-minute failover time typically required by service-level agreements (SLA).

Automated failover

Failover testing during this program utilized manual, not automated, procedures. No automation was used. Even though the manual process is quite simple, automation is available. The following tools provide automated failover:

- Storage System Scripting Utility (SSSU) allows the creation of scripts that can perform the functions outlined in "Failover testing results." SSSU allows you to tailor the failover procedures to your specific configuration.
- HP StorageWorks Cluster Extension EVA offers a majority node stretched cluster solution that would automatically fail over in the event of a disaster.

Failover best practices

Placing the remote Exchange server cluster in warm standby mode

Placing the remote Exchange server cluster in warm standby mode requires that the remote cluster be the active Exchange server. This requires presenting the Exchange Information Stores to the remote cluster. The procedure for warm standby is as follows.

Warm standby mode procedure		
Step	**Action**	**Details**
1	Make the remote Exchange cluster server the active Exchange server.	Follow the steps shown in "Planned failover procedure."
2	Put the remote Exchange cluster server in standby mode.	Using the Cluster Administrator, right-click the Exchange virtual server resource groups and select **Take Offline**. Wait for all resources to go offline. As an example, in Figure 31, right-click the **CAEXCHEVS1** and **CAEXCHEVS2** resource groups and select **Take Offline**.

3	Fail over the Information Store (storage) to the local site.	Using the CAUI:
		
		   a. Select the local EVA. As an example, in Figure 32, select **E54**.
		
		   b. In the Data Replication pull-down menu, select **Failover**.
		
		   c. In the Failover window (see Figure 33), click each DR group to be failed over, and then click **OK**. If all storage groups are to be selected, click **Select All**.
		
		A status window shows progress of the failover operation. In the Status window, click **OK** when done.

4	Ensure that the database/log disks are visible at the local site.	The cluster should see the disks automatically. If they do not become visible, use the Logical Disk Manager to rescan for the disks. Reboot if necessary.
5	Bring the local Exchange virtual server resource groups online.	Using the Cluster Administrator, right-click the Exchange virtual server resource groups and select **Bring Online**. Wait until all resources come online and the Information Stores are mounted. As an example, in Figure 31, right-click the **CAEXCHEVS1** and **CAEXCHEVS2** resource groups and select **Bring Online**.

Failover procedures

There are two types of failover procedures: planned and unplanned. Planned failovers are used to fail over to the remote site in non-disaster situations, for example, to perform regular maintenance or upgrades to the local Exchange server cluster. An unplanned failover is used in a disaster.

> **Note:** Not all possible failover scenarios were tested. Failover procedures must be tailored to each configuration.

Planned failover procedure

Planned failover procedure		
Step	Action	Details
1	Take the local Exchange server cluster offline.	Using the Cluster Administrator, right-click the Exchange virtual server resource groups and select **Take Offline**. Wait for all resources to go offline. As an example, in Figure 31, right-click the **CAEXCHEVS1** and **CAEXCHEVS2** resource groups and select **Take Offline**.

2	Fail over the Information Store (storage) to the remote site.	Using the CAUI:
		
		   a. Select the remote EVA. In Figure 32, select **E55**.
		
		   b. In the Data Replication pull-down menu, select **Failover**.
		
		   c. In the Failover window (see Figure 33), click each DR group to be failed over, and then click **OK**. If all storage groups are to be selected, click **Select All**.
		
		A status window shows progress of the failover operation. In the Status window, click **OK** when done.

3	Ensure that the database/log disks are visible on the remote Exchange cluster.	Windows should see the disks automatically. If they do not become visible, use the Logical Disk Manager to rescan for the disks. Reboot if necessary.
4	Bring the remote Exchange virtual server resource groups online.	Using the Cluster Administrator, right-click the Exchange virtual server resource groups and then select **Bring Online**. Wait until all resources come online and the Information Stores are mounted. As an example, in Figure 31, right-click the **CAEXCHEVS1** and **CAEXCHEVS2** resource groups and then select **Bring Online**.

Unplanned failover procedure

Regardless of the failure (loss, storage, or site), the failover procedure is very similar. The failover steps are as follows.

| Unplanned failover procedure		
Step	Action	Details
1	Fail over the Information Store (storage) to the remote site.	Using the CAUI: a. Select the remote EVA. In Figure 32, select **E55**. b. In the Data Replication pull-down menu, select **Failover**. c. In the Failover window (see Figure 33), click each DR group to be failed over, and then click **OK**. If all storage groups are to be selected, click **Select All**. A status window shows progress of the failover operation. In the Status window, click **OK** when done.
2	Ensure that the database/log disks are visible on the remote Exchange cluster.	Windows should see the disks automatically. If they do not become visible, use the Logical Disk Manager to rescan for the disks. Reboot if necessary.
3	Bring the remote Exchange virtual server resource groups online.	Using the Cluster Administrator, right-click the Exchange virtual server resource groups and then select **Bring Online**. Wait until all resources come online and the Information Stores are mounted. As an example, in Figure 31, right-click the **CAEXCHEVS1** and **CAEXCHEVS2** resource groups and then select **Bring Online**.

Figure 32. Continuous Access user interface
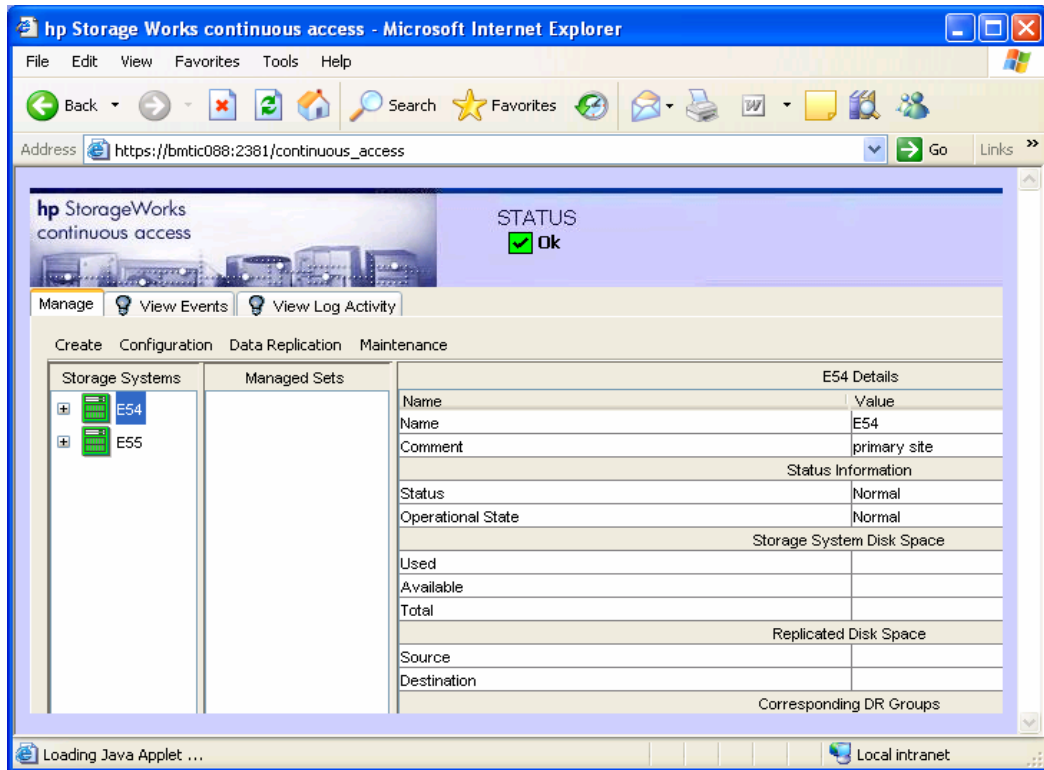
Figure 33. CAUI data replication failover window
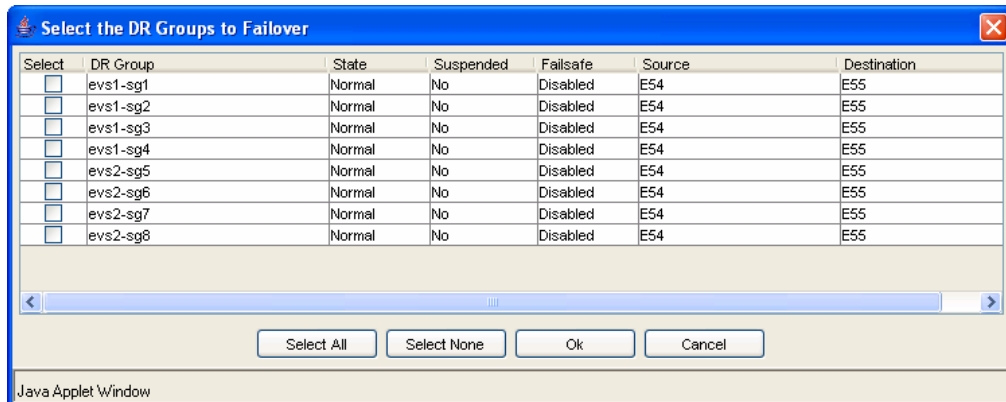
Continuous Access write history log merge testing

The third goal of this program was to determine the effect of a Continuous Access write history log (WHL) merge operation on Exchange performance. What is the effect of losing the ISL during normal Exchange operations? What does the added load of a Continuous Access log merge operation do to write latencies?

To determine the effect, the following test was performed:

1. Initiated a LoadSim 5,000-user MMB3 workload with a DR latency of 1.0 ms.
2. After two hours, disabled the ISL between the two sites, causing Continuous Access to log all changes (write history logging). During logging, the WHL for one DR group increased to 13 GB. The other storage groups were not monitored.
3. After two hours of logging, re-established the ISL, causing Continuous Access to merge the logged changes while Exchange traffic continued. The log merge operation took 50 minutes to complete for all storage groups. The 13-GB WHL was emptied at a rate of approximately 260 MB/min.

Results indicate that both the Exchange virtual servers and LoadSim clients were impacted by the log merge operation to a large degree. The log merge operation probably would have had a negative impact on Exchange client experience and response times. At a minimum, the testing showed that Exchange activity could continue during the ISL outage. It also showed that Exchange activity could continue during the log merge operation even with increased latencies. In a customer environment, a determination has to be made as to how best to handle a log merge situation. What level of degraded Exchange response times can be tolerated? It will differ from customer to customer. For possible options, see "Write history log merge alternatives" on page 53.

Write history log merge effect on Exchange servers

The Exchange virtual servers were affected in both their disk I/O performance and their ability to respond to client requests in a timely manner.

Figure 34 shows the read and write latencies over the entire operation. Clear delineations between the different operations are evident. The blue line represents write latencies. The first third of the graph shows write latencies when the ISL is present. The middle portion shows that the write performance improves when the ISL has been disabled, as all writes are local. The last portion shows greatly increased write latencies during the log merge operation after the ISL was re-established.

**Figure 34.** Write history log merge write latencies
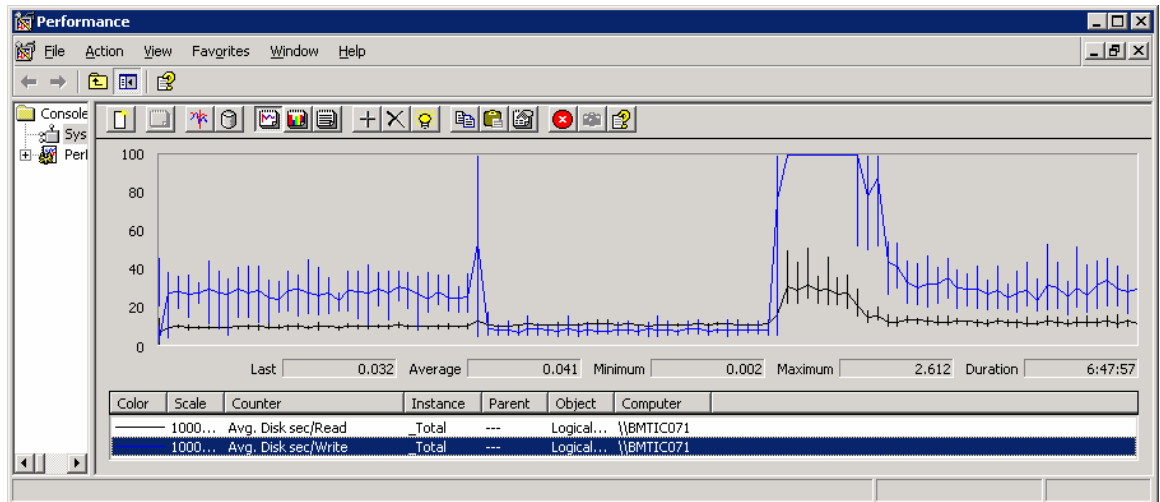
Looking just at the log merge operation, one can see that the write latencies average well above the desired 20-ms limit. The average was 171 ms and the maximum as high as 401 ms, as shown in Figure 35. When broken down, the log write latencies averaged 24 ms and spiked as high as 164 ms. Database write latencies averaged 189 ms and were as high as 724 ms.

**Figure 35.** Write history log merge disk latencies
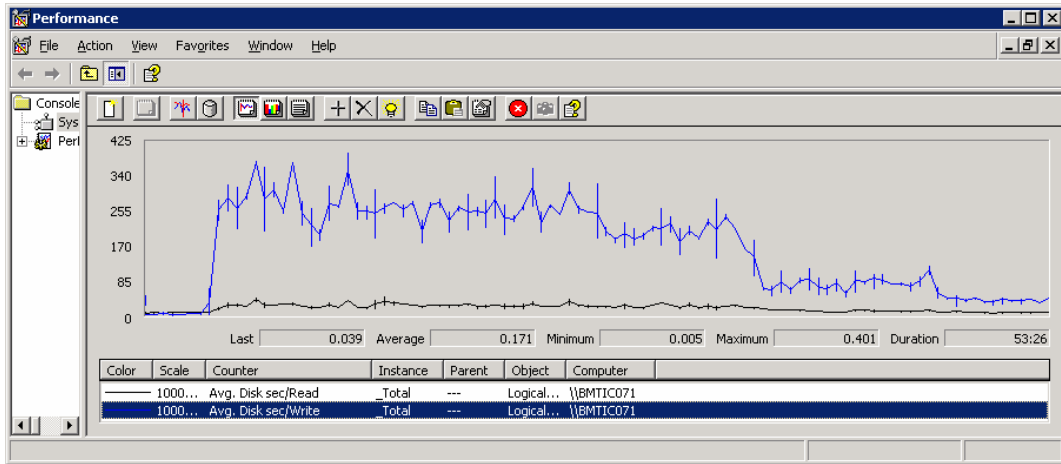
Average disk queue lengths were also greatly increased. The average was 142, while the maximum spiked as high as 266 (see Figure 36). Note that the scale of Figure 36 is set to 1000.

Figure 36. Write history log average disk queue lengths
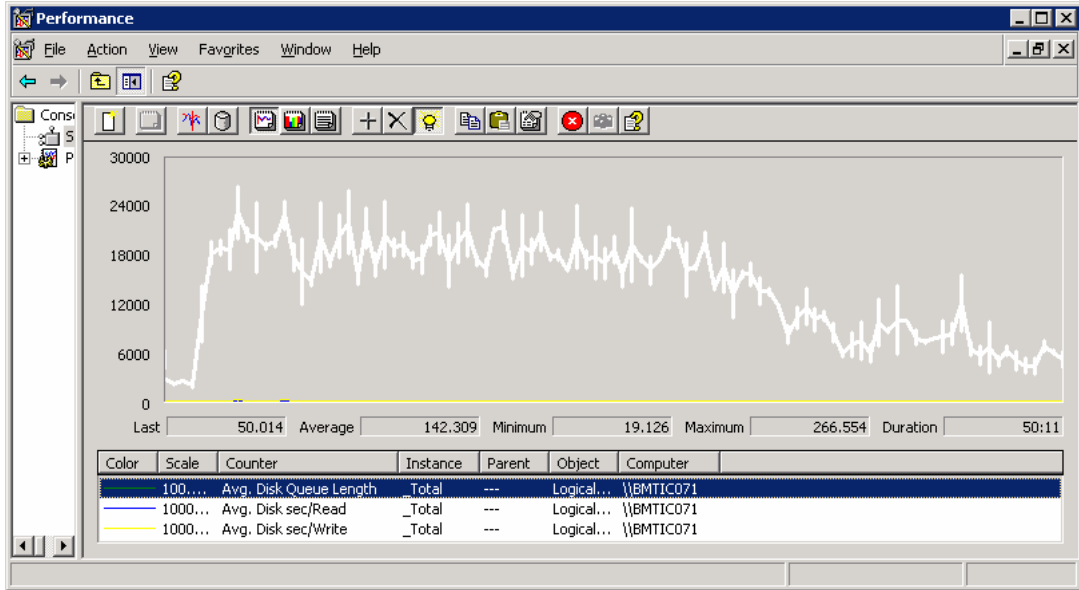
The Exchange servers' RPC request rate increased. Average RPC requests were 20 but increased as high as 51. Average RPC latency was 52 but spiked as high as 204. The average experienced is still very close to acceptable RPC latency limits. Figure 37 shows RPC values during the log merge operation. The RPC request rate returned to a steady state after the 50-minute log merge operation.

Figure 37. Average RPC requests and latencies
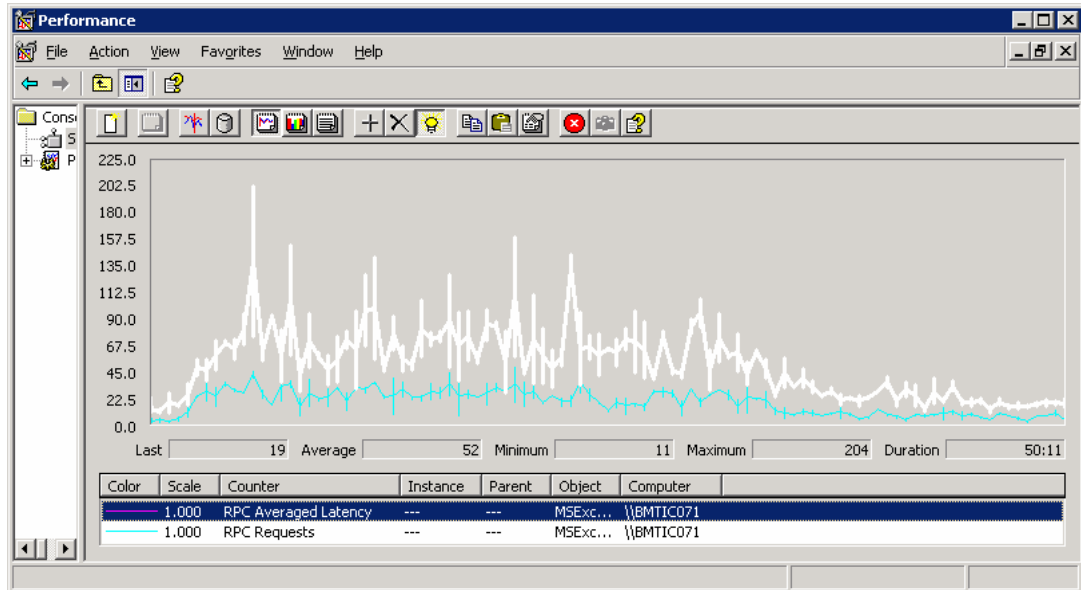
The data clearly shows that the added load of the log merge operation affected the ability of Exchange to respond effectively to client requests.

Write history log merge effect on LoadSim clients

From the client perspective, the log merge operation caused clients to wait much longer for their requests to be processed. Figure 38 shows this effect during the log merge operation. The clients saw request latencies that averaged 405 ms. Before the loss of the ISL, the client request latency averaged 105 ms. The merge operation caused a fourfold increase in request latency.

Figure 39 shows the CPU usage and request latency for another client. The green line shows that CPU usage increased to 100% for the period of the loss of the link and the log merge operation. This behavior was seen on other client servers as well.

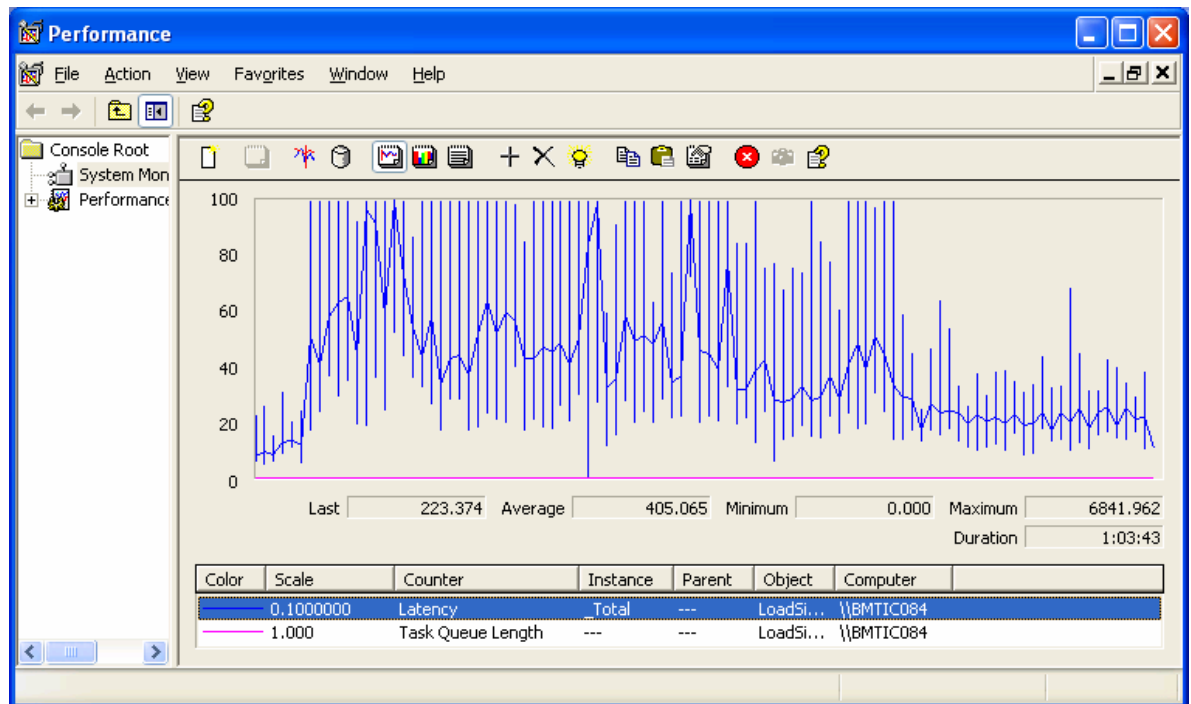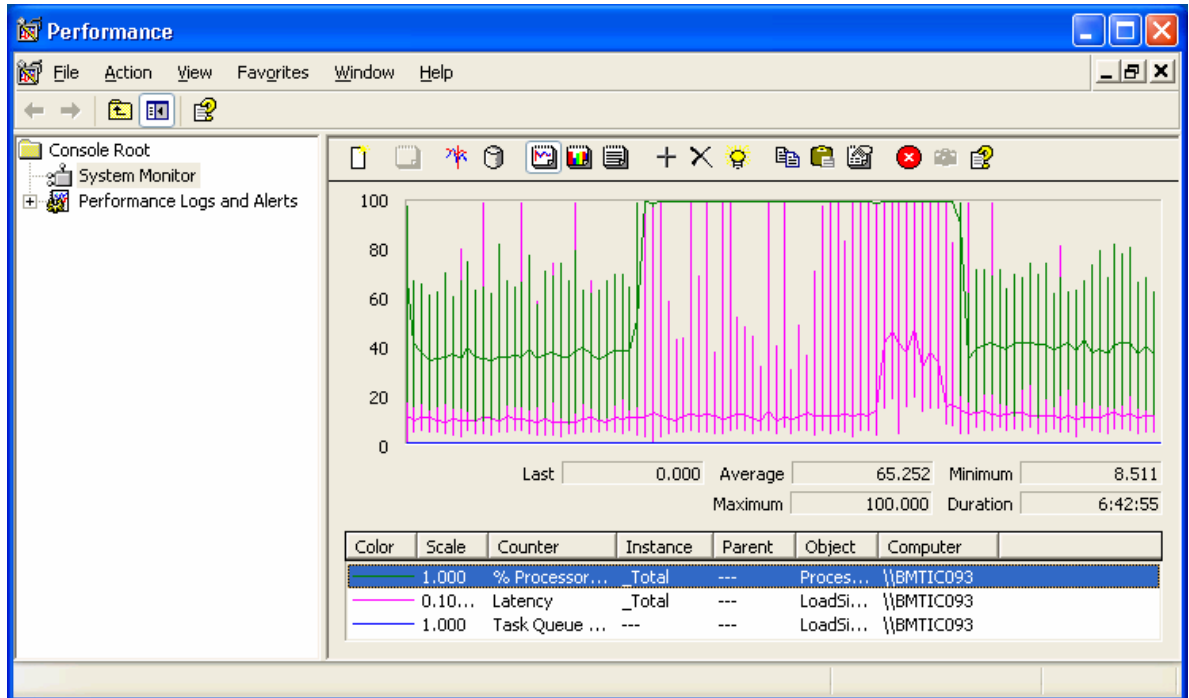**Figure 38.** LoadSim client BMTIC084 latency

Figure 39. LoadSim client BMTIC093 CPU usage and latency

Write history log merge alternatives

Several alternatives exist for the log merge operation. It has been shown in this testing that the log merge will affect Exchange's ability to process email requests in a timely fashion. Following are choices for handling the log merge:

- Stagger the log merge operation across the DR groups instead of merging all storage groups simultaneously. The administrator would suspend all DR groups and resume one at a time until all DR groups have been normalized. Normal Exchange activity would continue during this operation.
- Delay the merge operation to hours of low Exchange activity. This is an option if sufficient disk space for write logging exists. All DR groups would be suspended when the ISL is re-established. The user can perform the log merge operation during off hours. With VCS 3.020 and later, a fast re-sync option is available that may reduce the time to perform the full copy data.

> **Note:** This option may greatly increase disk space requirements and the time to perform the merge operation, especially if the ISL loss occurs early in the day. Users would have to determine their write logging capacity based on an outage during peak hours. It is imperative that users monitor write log disk usage because when all available write history log space is used, Continuous Access suspends write history logging and marks the DR groups for a full copy operation.

- Request users to either suspend or reduce Exchange usage during log merge operations.
- Do nothing. Customers could choose to ride out a period of lower client response times.

Further testing

This test exercise has brought up areas that could use further exploration:

- Log-only replication
- Replication over lower speed links such as T3 or OC3
- HP Cluster Extension EVA (geographically dispersed clusters based on Microsoft Cluster Server)

Appendix A—Configuration bill of materials (BOM)

Table A-1. Configuration BOM

	Quantity	Part number
Exchange Server 2003 clusters		
HP ProLiant BL20p G2 server X3.06-512KB, 1GB , 2 processor, FC mezzanine	6	323148-B22
HP BL p Class Server Blade Enclosure	1	281404-B21
HP BL p Class RJ-45 Patch Panel	1	230766-B21
4-GB PC2100 DDR SDRAM DIMM 4x1024WW	6	300680-B21
72-GB 10K U320 Pluggable Hard Drive	12 (2 per server)	286714-B22
HP StorageWorks Secure Path v4.0C SP1 for Windows, 1 license	1 per server	165989-B24
Microsoft Windows Server 2003 Enterprise Edition	1 per server	
Microsoft Exchange 2003 Enterprise Server SP1	1 per server	
Exchange clients		
HP ProLiant DL320 G2 server, Pentium® 4, 2.7 GHz, 640 MB	8	
Microsoft Windows XP Professional SP2	1 per client	
Microsoft Exchange Server Load Simulator 2003	1 per client	
Microsoft Office Outlook 2003	1 per client	
General-purpose server (Management Appliance)		
HP ProLiant DL320 G2 server, Pentium 4, 2.7 GHz, 640 MB	1 per site	
Microsoft Windows Server 2003 Enterprise Edition	1 per site	
HP StorageWorks Continuous Access V1.2	1per server	
HP StorageWorks Command View EVA V3.3	1 per server	
Storage—EVA5000		
HP StorageWorks EVA 2C12D—168 drives with 72-GB 10K drives	1 per site	283199-B23
HP StorageWorks Virtual Controller Software Package VCS v3.020 for Dual HSV110 Controllers (VCS v3.0)	1 per array	250203-B25
HP StorageWorks VCS V3.0 Platform Kit for EVA5000—Windows 2003	1 per array	333685-B21
Fibre Channel cables	As required	221692-B2x
HP StorageWorks Continuous Access EVA media and documentation CD v1.2 user interface	1 per array	331268-B22
HP StorageWorks Continuous Access EVA5000 6-TB license v1.0	1 per array	331273-B21

(Table A-1 continued next page)

Table A-1. Configuration BOM *(continued)*

	Quantity	Part number
LAN/SAN infrastructure		
HP ProCurve 2708 Gigabit Ethernet switch	2	J4898A
HP StorageWorks SAN Switch 2/16 Brocade 3800	2 (1 per site)	322118-B21
HP Short Wave Optical Transceivers	16 per switch	221470-B21
HP StorageWorks SR2122-2 FCIP gateways	2 (1 per site)	350074-B21
HP 10-km Long-Distance Transceiver Kit, 2 Gbps, LC (All SAN Switch 2 models)	1 per Brocade 3800	300835-B21

Appendix B—Exchange storage configuration (Phase 1)

Table B-1. EVA storage group configuration

Disk group	Disks	Capacity	Contents
Disk Group 1	36	2304 GB	EVS1 Databases
Disk Group 2	36	2304 GB	EVS2 Databases
Disk Group 3	16	948 GB	EVS ½ Logs
Disk Group 4	52	3388 GB	Write history logging, backup storage

Table B-2. LUN mapping

Disk drive letter	Data type	Disk group	LUN size (GH)	RAID level	LUN
Q$	Quorum	DG3	10	1	1
E:\Exchsrvr (mount point)	SMTP	DG1	50	1	2
E$	SG1-Data	DG1	125	1	3
F$	SG2-Data	DG1	125	1	4
G$	SG3-Data	DG1	125	1	5
H$	SG4-Data	DG1	125	1	6
E:\Log (mount point)	SG1-Log	DG3	50	1	7
F:\Log (mount point)	SG2-Log	DG3	50	1	8
G:\Log (mount point)	SG3-Log	DG3	50	1	9
H:\Log (mount point)	SG4-Log	DG3	50	1	10
I:\Exchsrvr (mount point)		DG2	50	1	15
I$	SG5-Data	DG2	125	1	16
J$	SG6-Data	DG2	125	1	17
K$	SG7-Data	DG2	125	1	18
L$	SG8-Data	DG2	125	1	19
I:\Log (mount point)	SG5-Log	DG3	50	1	20
J:\Log (mount point	SG6-Log	DG3	50	1	21
K:\Log (mount point)	SG7-Log	DG3	50	1	22
L:\Log (mount point)	SG8-Log	DG3	50	1	23

Table B-3. EVS data replication (DR) groups

DR group	Member
EVS1-SG1	EVS1-SMTP
	SG1-Data
	SG1-Log
EVS1-SG2	SG2-Data
	SG2-Log
EVS1-SG3	SG3-Data
	SG3-Log
EVS1-SG4	SG4-Data
	SG5-Log
EVS2-SG5	EVS2-SMTP
	SG5-Data
	SG5-Log
EVS2-SG6	SG6-Data
	SG6-Log
EVS2-SG7	SG7-Data
	SG7-Log
EVS2-SG8	SG8-Data
	SG8-Log

Appendix C—Microsoft Load Simulator performance guidelines

The information in Table C-1 was taken from *Using Microsoft Exchange Server Load Simulator 2003*. The table shows the performance counters collected during LoadSim testing. It also shows recommended performance values.

Table C-1. Microsoft Load Simulator performance guidelines

Object and counter	Description	Expected result
Processor: % Processor Time (_Total)	Displays the total processor time.	As the load increases, this counter increases to almost 100%, at which point the processor is bottlenecked, and no additional rate increases are possible.
MSExchangeIS: RPC Requests	Displays the number of incoming RPC requests from clients.	This value should be below 30 at all times.
MSExchangeIS: RPC Averaged Latency	Displays the average latency of the RPC requests.	This value should be below 100 ms at all times.
MSExchangeIS: VM Largest Block Size	Displays the largest virtual memory block size available in the Exchange store.	This value should be above 500 MB at all times.
MSExchangeIS: VM Total 16MB Free Blocks	Displays the number of 16-MB free blocks available in the Exchange store.	This value should be above 1 at all times.
Processor: % Processor Time (_Total)	Displays the total processor time.	As the load increases, this counter increases to almost 100%, at which point the processor is bottlenecked, and no additional rate increases are possible.
MSExchangeIS: VM Total Free Blocks	Displays the number of free blocks available in the Exchange store.	This value should be above 1 at all times.
MSExchangeIS: VM Total Free Blocks Bytes	Displays the total size of all free blocks available in the Exchange store.	This value should be above 500 MB at all times.
Memory: Available Mbytes	Displays the available memory in MB on the server.	This value should be above 500 MB at all times.
Memory: Pool Nonpaged Bytes	Displays the total non-paged bytes (kernel memory) used in the system.	This value should be below 80 MB at all times.
Memory: Pool Paged Bytes	Displays the total paged bytes (kernel memory) used in the system.	This value should be below 200 MB at all times.

(Table C-1 continued next page)

Table C-1. Microsoft Load Simulator performance guidelines *(continued)*

Object and counter	Description	Expected result
Memory: Free System Page Table Entries	Displays the total number of PTEs in use in the system. If this number reaches zero, the system will be in an unstable state.	This value should never reach zero.
Physical disk: Average Disk Queue Length	Displays the average disk queue length.	The number for all drives should reach zero every minute or less.
Physical disk: Current Disk Queue Length	Displays the current disk queue.	The number for all drives should reach zero every minute or less.
Physical disk: Average Sec/Read and Average Sec/Write	Displays the disk latency.	The latency for all drives should be below 100 ms at all times.
MSExchangeIS Mailbox: Send Queue Size	Displays the number of items awaiting mail delivery, that is, the number of items waiting to move from the mailbox store to the SMTP queue in Internet Information Services (IIS) for categorization. SMTP determines the appropriate server to which to send mail.	The average number should be less than 1% of the number of users in the simulation. Average queue size should not increase during the test. It should remain relatively constant.
MSExchangeIS Mailbox: Receive Queue Size	Displays the number of items waiting to move from the SMTP queue to the mailbox store.	The average number should be less than 1% of the number of users in the simulation. Average queue size should not increase during the test. It should remain relatively constant.
SMTP server: Categorizer queue length	Displays the queue of items waiting to be categorized in IIS.	Average queue size should not increase during the test. It should remain relatively constant. It should be less than 10.
SMTP server: local queue length	Displays the queue of items waiting in IIS to be delivered to the local mailbox store after categorization.	Average queue size should not increase during the test. It should remain relatively constant. It should be less than 1,000 throughout the test.
SMTP server: remote queue length	Displays the queue of items in IIS waiting to be delivered on another server after categorization.	Average queue size should not increase during the test. It should be less than 1,000 throughout the test.
SMTP server: NDRs Generated	Displays the number of Non-Delivery Receipts sent.	This value should be zero during the entire test.

(Table C-1 continued next page)

Table C-1. Microsoft Load Simulator performance guidelines *(continued)*

Object and counter	Description	Expected result
LoadSim global: task queue length	Displays the queue of pending tasks on a LoadSim client.	Average queue size should not increase during the test. It should be less than 100 throughout the test.
LoadSim action: latency (_total)	Indicates LoadSim client response times.	This value must not exceed 1,000 ms for any of the LoadSim clients used on the server.
MSExchangeIS mailbox: messages submitted per minute	Indicates the message submission rate.	System Monitor data should match the LoadSim predicted value for message submission rate. Messages submitted for each MMB3 user in an eight-hour test is 85. The expected average rate is: Messages submitted per minute = (0.1771) * (number of users) **Note:** This rate varies during a test run because of users logging on and off and users replying to and forwarding messages.
MSExchangeIS mailbox: message recipients delivered per minute	Indicates the message delivery rate.	System Monitor data should match the LoadSim predicted value for message received rate. Messages received per MMB3 user over an eight-hour period are 210. The expected average rate is: Message recipients delivered per minute = (0.4375) * (number of users)
MSExchangeIS Transport Driver: TempTable Current	This is the queue of items waiting to go from the Exchange store to IIS.	Average queue size should not increase during the test. It should be less than 1,000 throughout the test.

Appendix D—Reference documentation

The following documentation was used during this testing process:

HP documentation

- HP Best Practices for Microsoft Exchange Server 2000 and 2003 Cluster Deployments
  http://h71028.www7.hp.com/enterprise/cache/70586-0-0-225-121.aspx
- HP StorageWorks Continuous Access Planning Guide—AA-RW1DA-TE
  http://h18006.www1.hp.com/products/storage/software/conaccesseva/index.html
    Click **Technical Documentation.**
- HP StorageWorks Continuous Access EVA user guide performance estimator—AA-RU5ME-TE
  http://h18006.www1.hp.com/products/storage/software/conaccesseva/index.html
    Click **Related information** under the Product Information tab.

Microsoft documentation

- Deployment Guidelines for Exchange Server Multi-Site Data Replication
  http://www.microsoft.com/technet/prodtechnol/exchange/guides/E2k3DataRepl/bedf62a9-dff7-49a8-bd27-b2f1c46d5651.mspx
- Multi-site data replication support for Exchange 2003 and Exchange 2000
  http://support.microsoft.com/?kbid=895847
- Optimizing Storage for Exchange Server 2003
  http://www.microsoft.com/technet/prodtechnol/exchange/2003/library/optimizestorage.mspx
- Exchange Server 2003 Technical Reference Guide
  http://www.microsoft.com/technet/prodtechnol/exchange/guides/E2k3TechRef/4c8062cb-f07d-463b-a297-00614095ca24.mspx
- Troubleshooting Microsoft Exchange Server 2003 Performance
  http://www.microsoft.com/technet/prodtechnol/exchange/guides/TrblshtE2k3Perf/b22501d4-80cd-4038-b2bb-b4672cd68870.mspx
- Exchange Server Load Simulator 2003
  http://www.microsoft.com/downloads/details.aspx?FamilyID=92eb2edc-3433-47ca-a5f8-0483c7ddea85&DisplayLang=en
- Exchange Server 2003 Jetstress Tool
  http://www.microsoft.com/downloads/details.aspx?familyid=94b9810b-670e-433a-b5ef-b47054595e9c&displaylang=en
- Microsoft Exchange Server Best Practices Analyzer Tool
  http://www.microsoft.com/exchange/downloads/2003/ExBPA/default.mspx

For more information

- HP Storage Solutions for Microsoft Exchange Server 2003:
  http://www.hp.com/solutions/microsoft/exchange/storage
- Microsoft Exchange Server: http://www.microsoft.com/exchange
- HP StorageWorks Continuous Access:
  http://h18006.www1.hp.com/products/storage/software/conaccesseva/index.html
- HP SWD Customer Focused Testing e-mail: SWD_WW_CFT_TEAM@hp.com