WHITE PAPER

August 1997

Prepared By
Internet Solutions
Business Unit

Compaq Computer
Corporation

CONTENTS

Performance of Microsoft Exchange Server 5.0 on Compaq ProLiant 6000-Class Servers

Electronic Messaging and GroupWare applications are quickly becoming ubiquitous in the business world. Because people must process and share increasing amounts of information, there is a need for powerful software and hardware products to satisfy the requirements of these applications. As customers deploy such systems, capacity planning and optimization become increasingly critical. Being the world's largest supplier of Intel-based servers, Compaq is uniquely positioned to be a leader in GroupWare and Messaging system platforms.

Microsoft Exchange Server has been the focal point of extensive development and testing by both Microsoft and Compaq. Throughout this activity, Compaq and Microsoft have worked together to optimize Microsoft Exchange Server performance on Compaq Server products.

Microsoft Exchange Server is complex and has many features and capabilities. For system administrators as well as potential purchasers of both Microsoft Exchange Server and Compaq products, a solid grasp of the performance issues associated with Microsoft Exchange Server is crucial to the decision-making process.

This white paper contributes to such an understanding by examining various performance aspects of Microsoft Exchange Server 5.0 on Compaq ProLiant Servers, specifically the ProLiant 6000-class of servers. The reader will learn what to expect from such a system under various user loads. Also, he or she will discover other useful information pertaining to Microsoft Exchange Server configuration and performance tuning, as well as CPU, RAM, and disk utilization on Compaq ProLiant servers. Ultimately, this paper should help the potential purchaser make good decisions about which Compaq products to purchase in order to satisfy the requirements of his or her business needs.

COMPAQ

NOTICE

Performance of Microsoft Exchange Server 5.0 on Compaq ProLiant 6000-Class Servers

First Edition (August 1997)
ECG052.0897

EXECUTIVE SUMMARY

The Compaq ProLiant 6000 is the flagship of Compaq's server product line. This server is an ideal deployment platform for Microsoft Exchange. With the current version of Microsoft Exchange Server (version 5.0), the ProLiant 6000-class system with two processors and 256 MB of RAM can easily accommodate user loads of up to 3000 users (**Medium** profile). A ProLiant 6000-class system with one processor and 128 MB of RAM can support up to 2000 users. Future versions of Exchange Server will allow this platform to scale beyond the 3000-user limit and maximize the full power of the Compaq system architecture.

Compaq tested the ProLiant 6000 in two hardware configurations at various user loads (see *Test Configuration* below). Configuration A was a single-processor system with 128 MB of RAM and an entry-level disk subsystem. Configuration B was a two-processor system with 256 MB of RAM and a more scalable disk subsystem. Before encountering limits in Exchange Server, both configurations showed excellent scalability at various user regressions (see *Performance Results*).

As the leader in the Intel-based server market, Compaq is uniquely positioned to be the server of choice for deployment of GroupWare and Messaging systems such as Microsoft Exchange Server. The entire Compaq server product line offers scalable server platforms positioned at key price points offering differentiated feature sets. From the ProSignia 200 and ProLiant 800/850R to the ProLiant 2500 through the ProLiant 6000 and beyond, Compaq servers represent the ultimate combination of price/performance, manageability, integration, and ease of use.

When deploying messaging systems like Microsoft Exchange Server, one must consider not only the performance and capacity of the server but also the price/performance. Several competing hardware vendors offer RISC-based platforms capable of supporting equivalent user loads, but provide these systems at significantly higher cost. Only Compaq delivers industry-leading performance combined with lowest total cost of ownership. Another important consideration is the mission-critical nature of GroupWare and messaging within your organization. It is one thing to be able to support 3000 users on a single server, and another thing to recover from the loss of productivity when this single potential point of failure is unavailable. The question is not whether you can support thousands of users on a single system, but whether you want to.

Compaq has invested heavily in GroupWare and messaging applications in several areas. First, applications like Microsoft Exchange Server have been optimized and integrated for Compaq servers and provided on Compaq SmartStart. SmartStart allows users of Exchange Server to integrate with confidence by facilitating the installation of Exchange Server on Compaq servers without errors. Another area of significant investment is the Intranet/GroupWare Engineering group at Compaq. This group supports applications like Exchange Server and ensures their optimal performance and integration on Compaq servers. Compaq has engineering resources working on site at Microsoft with various BackOffice groups like the Exchange team to ensure tightly coupled integration and superior performance on Compaq servers.

TEST CONFIGURATION

Introduction

Two main tools were used to prepare the data for this white paper:

- Microsoft Exchange Server LoadSim 5.0
- Microsoft Windows NT Performance Monitor

There is an important point to remember when reviewing the data and assumptions in this white paper:

A simulation is only as good as the workload characterization.

The results presented in this paper may or may not represent your production messaging environment. Due to the complexity of Microsoft Exchange deployments and the number of options with which to deploy the product, careful study and characterization of your messaging environment should be done. To determine the actual performance of a Compaq server product in your environment, it should be characterized, profiled, and simulated against your planned deployment platform. This paper is meant to serve as a tool for making assumptions about Exchange Server performance under different user loads and Compaq hardware configurations.

The configuration of test runs selected for this paper is based on a LoadSim **Medium** profile. This profile most closely represents a typical corporate mail user. When conducting your own Exchange capacity planning studies, it is important to characterize the workload of the mail users in your environment. This can be accomplished in a variety of ways with tools available from both Microsoft and third parties. For more information regarding LoadSim, contact Microsoft Product Support Services or Microsoft on the Web at **http://www.microsoft.com**.

User-initiated Actions and “The Quantum Effect”

One way to think of a user-initiated action is as a single, inseparable “quantum” event. In this context a quantum event is one that cannot be subdivided into smaller, component events. Such events have also been called “atomic.”

Even though user-initiated actions such as reading a message invoke several sequential operations on the client and server computers, they can be treated as quantum events for the purpose of load analysis. In this case, the load a user places on a server can be regarded as evenly distributed in time only when the number of users on the server is high enough to average out the quantum effects of individual usage patterns.

For example, in any system there will be bursts of activity due to coincidence, when more users than usual initiate actions at nearly the same time. These bursts can be random in nature or a consequence of scheduled events that affect many users on the same server, such as when a large meeting adjourns. Additionally, the near-real-time notification mechanisms built into Microsoft Exchange Server may contribute to the quantum effect. Users on the same server are notified that a new message has arrived at essentially the same time. If the message was sent to a large number of users, a burst of activity may occur as many of those users read the message immediately upon receipt.

Fortunately, the quantum effect becomes less and less significant as the number of users hosted per server increases. Because of the large number of users Microsoft Exchange Server can typically host, even on small servers, only the largest variations in user activity are typically noticeable by the server. From the user’s perspective, however, the quantum effect is still noticeable if the server is under load. The sudden increase in server load affects the delay between the initiation of an action by the user and the server’s response to it.

LoadSim

The main tool used in generating the performance data contained in this paper was the Microsoft Exchange Server User Load Simulation utility called LoadSim. As its name implies, LoadSim is a tool for simulating a client user load on an Exchange Server. Its purpose is to enable a single Windows NT machine—called a LoadSim client—to simulate multiple Microsoft Exchange client users. The current version of LoadSim not only provides simulation of MAPI protocol clients but also provides support for POP3, NNTP, LDAP, and HTTP client protocols.

The operation of LoadSim users is governed by a LoadSim profile. This profile controls factors such as how long a LoadSim ‘day’ is, how many email messages to send in a day’s time, how many times to open and read existing email, whether to use distribution lists, whether to use public folders, etc.

LoadSim is a simulator and therefore not a perfect image of real-world activity. Because of the nature of the product and the quantum effect discussed previously, it is impossible to fully emulate a client. Thus there are two points to remember when considering this data: First, the LoadSim results do not properly address users’ logging on and off. When a test is run, all users log on sequentially. Once this process is complete, users begin their tasks. If users typically log on and off multiple times during the day, the NT logon can have an impact on server and network rates of utilization, especially when many users log on simultaneously. Based on your organization’s implementation, these factors could pose important considerations in addition to those presented in this paper. The second point to remember is that the response numbers generated by LoadSim are based on server response to clients and do not account for strictly client-side actions such as rendering rich text once a message is received.

LoadSim, despite possible shortcomings, creates a highly accurate simulation of reality. It mimics the full Microsoft Exchange Client in many respects. First, it uses .MSG files, the same format used by the Exchange Client. This guarantees that messages generated by LoadSim have the same properties as those sent by real users of the Exchange Client. Second, LoadSim uses the same remote procedure call (RPC) semantics as those used by the Client. Third, LoadSim registers MAPI change notifications in the same manner as they are registered by the Client. Finally, LoadSim even emulates the Microsoft Exchange Client list-box cache, which the Client uses for folder and message panes in the viewer when a user browses and selects messages on the server.

Default User Profiles

There are three pre-configured profiles built into LoadSim: **Light**, **Medium**, and **Heavy**. Their characteristics are detailed in Table 1.

DEFAULT LOADSIM USER PROFILES

LoadSim USER ATTRIBUTE	ATTRIBUTE DETAIL	LIGHT	MEDIUM	HEAVY
TEST DURATION	Length of a day (hours)	8	8	8
READING MAIL	New mail (times/day)	12	12	12
	Existing mail (times/day)	5	15	20
AFTER READING MAIL	% of Reply	5%	7%	15%
	% of Reply All	3%	5%	7%
	% of Forward	5%	7%	7%
	% of Move	20%	20%	20%
	% of Copy	0%	0%	0%
	% of Delete	40%	40%	40%
	% of Do nothing	27%	21%	11%
DISTRIBUTION LISTS	Minimum size	4	4	4
	Maximum size	50	50	50
	Average size	10	10	10
	Cover 100% of users (no overlap)	Yes	Yes	Yes
ATTACHMENTS	% to Run/Load Mail Attachment (if one exists)	25%	25%	25%
INBOX SIZE	Inbox Size Limit (# messages)	20	125	250
SENDING MAIL	New mail (times/day)	2	4	6
	Save a copy in Sent Mail Folder?	Yes	Yes	Yes
	Number of random recipients	3	3	3
	% of time to add a Distribution List	30%	30%	30%
	Message Priority	Normal	Normal	Normal
	Delivery Receipt?	No	No	No
	Read Receipt?	No	No	No
NEW MAIL MESSAGE CONTENT *Text-only, no attachment*	1K body (ups1K.msg)	90%	64%	50%
	2K body (ups2K.msg)	0%	17%	10%
	4K body (ups4K.msg)	0%	4%	5%
NEW MAIL MESSAGE CONTENT *1K mail body, with attachment*	10K attachment (ups10Kat.msg)	10%	5%	10%
	Embedded bitmap object (upsBMobj.msg)	0%	2%	5%

LoadSim USER ATTRIBUTE	ATTRIBUTE DETAIL	LIGHT	MEDIUM	HEAVY
	Word attachment (upsWDatt.msg)	0%	2%	5%
	Excel attachment (upsXLatt.msg)	0%	4%	5%
	Embedded Excel object (upsXLobj.msg)	0%	2%	10%
SCHEDULE+ CHANGES	Changes per day	1	5	10
	Update Free/Busy information?	No	No	No
	Average Schedule File Size	22K	22K	22K
PUBLIC FOLDERS	Folder activity	None	None	None
CALCULATED DAILY LOAD (based on these defaults)	TOTAL MAIL RECEIVED PER DAY	22.94	66.30	118.89
CALCULATED DAILY LOAD (based on these defaults)	TOTAL MAIL SENT PER DAY	4.70	14.18	30.67
	Mail sent as New mail	2.00	4.00	6.00
	Mail sent as a Reply	1.05	3.76	13.03
	Mail sent as a Reply to All	0.60	2.67	5.82
	Mail sent as a Forward	1.05	3.76	5.82
CALCULATED DAILY LOAD (based on these defaults)	AVERAGE # RECIPIENTS FOR EACH MESSAGE	4.88	4.68	3.88

Table 1. Default LoadSim user profile definitions

LoadSim Score

If we have several recorded response times for some action, why not just use the average response time, or the maximum, to determine acceptability? This approach does not work for several reasons.

The average response time does not tell you anything about the distribution of response times. For example, you will get the same average if all your actions took two seconds, or if half took one second and the other half took three seconds. However, there is a big difference in user perception between these two distributions.

We do not use the maximum response time here because of the quantum effect, explained in a previous section. In a user-driven client/server system like Microsoft Exchange Server, there is always a statistical chance that several clients will require the same resource at the same time, and some will have to wait longer than usual. As a result, the maximum response time is not really a fair measure of acceptability.

Instead, we use the 95th percentile (as recommended by Microsoft). If the 95th percentile response time for a set of actions is one second, that means 95 percent of the response times are at or below one second. Only five percent (one in twenty) of the response times exceeded one second. For comparison, the maximum response time is just the 100th percentile (100 percent of the response times are at or below the maximum). The **median** of a set of response times is defined as the 50th percentile (not to be confused with the mean, which is the average). The 95th percentile is statistically accurate and based on the entire set of response times, but is also fair to the quantum effect and real-world user perception.

The data point resulting from a LoadSim run is called the *Score*. The LoadSim Score represents a weighted average of the 95th percentile Exchange client response time (in milliseconds) for the various Exchange tasks. The READ task is the task weighted highest, accounting for over half of the score. The reason for this is that different client tasks (READ, BROWSE, DELETE, SEND, etc.) are perceived differently by the user. For example, a user will most likely be willing to wait two seconds for a SEND task but would not be willing to wait two seconds for a READ task.

Only you can determine the acceptable response time in your environment, but most capacity planners default to the Microsoft-recommended maximum for the MAPI protocol, which is sub-second response time (<1000 ms). One thousand milliseconds has been assumed as a safe upper limit for this paper. A lower Score indicates better Exchange Server performance.

NT Performance Monitor

TIP: Performance Monitor objects for several key Compaq server subsystems are available via the Web at **http://www.compaq.com** *or on the Compaq ResourcePaq for NT version 2.5*

The main tool used in monitoring and collecting the performance data contained in this paper is the Windows NT Performance Monitor (PerfMon).

PerfMon monitors performance objects and counters within Windows NT, and it is these objects and counters that depict how the Exchange Server machine is performing under load. Exchange Server, along with several Compaq server subsystems, has counters available for instrumented subsystems. Performance Monitor, although not perfect, is an invaluable tool for capacity planning and performance tuning for NT Server and applications. Best of all – it's free! Many similar tools in the Unix environment would cost thousands of dollars in addition to the cost of the operating system.

NOTE: Be sure to execute the NT command "DISKPERF –y" in order to monitor disk objects in NT Performance Monitor.

For more information on how to use PerfMon, refer to the Windows NT Resource Kit for Windows NT 4.0. The resource kit contains a thorough treatment of PerfMon and some good suggestions for detecting bottlenecks. The principles outlined there are relevant to monitoring the performance of Exchange Server. The Exchange Server Resource Kit contains complete details on Exchange-specific Performance Monitor objects.

Configuration of Test Facility

The test facility is configured on two 100BaseTX Fast Ethernet networks -- one network for the LoadSim clients and one network for data collection. This is to isolate the network traffic imposed by data collection from the actual test network traffic. Figure 1 shows the network topology.

There are twelve LoadSim client machines. As stated earlier, a LoadSim client is a Windows NT machine configured with the Microsoft Exchange client software and LoadSim. A single LoadSim client can simulate multiple users. In this case, each of the twelve LoadSim clients can simulate up to 500 users. The load imposed by each user is based on the **Medium** profile outlined in Table 1.

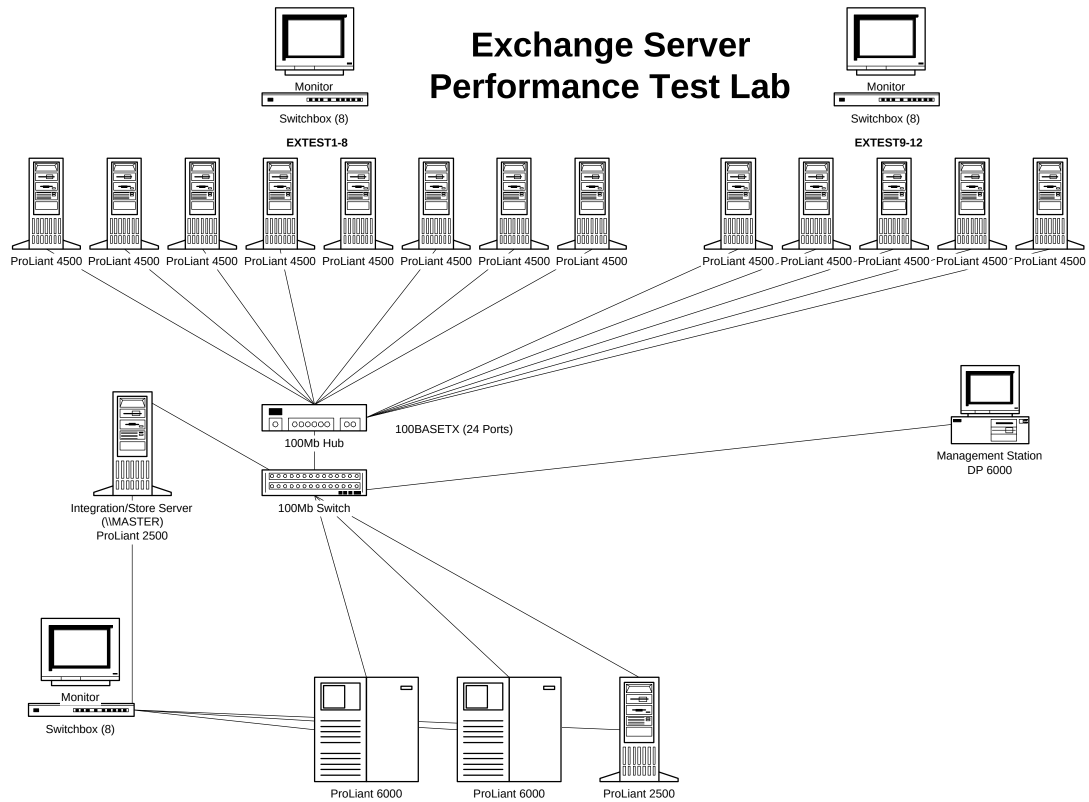

Figure 1. Exchange Server test lab setup and topology.

LoadSim Client Configurations

LOADSIM CLIENT HARDWARE CONFIGURATIONS

Machine Class	Compaq ProLiant 4500
System Processor	Dual Pentium/133 – 2-MB Cache
System Memory	128 MB
Disk Subsystem	4-GB Fast SCSI-2
Network Interface	Compaq NetFlex-3 (100BaseTX)
Operating System	Windows NT Workstation v4.0 + SP2

Table 2. LoadSim client hardware configurations

ProLiant 6000 Exchange Server Configurations

Results presented in this paper were based on two different ProLiant 6000 hardware configurations. The configurations were designed to provide for the availability of two distinct levels of hardware resource for Exchange Server. This gave the ability to compare performance characteristics under the same user loads with different levels of hardware resource. The tested configurations are detailed below.

PROLIANT 6000 HARDWARE CONFIGURATION A

Machine Class	Compaq ProLiant 6000
System Processor	(1) Pentium Pro/200 – 512K Cache
System Memory	128 MB
Disk Subsystem	(1) SMART-2/P OS/LOG/Pagefile: (2) 4.3-GB Drives – RAID1 IS: (7) 4.3-GB Drives – RAID5
Network Interface	Compaq NetFlex-3 (100BaseTX)
Operating System	Windows NT Server v4.0 + SP2 Exchange Server 5.0

PROLIANT 6000 HARDWARE CONFIGURATION B

Machine Class	Compaq ProLiant 6000
System Processor	(2) Pentium Pro/200 – 512K Cache
System Memory	256 MB
Disk Subsystem	(2) SMART-2/P OS/Pagefile: (2) 4.3-GB Drives – RAID1 LOG: (2) 4.3-GB Drives – RAID1 IS: (12) 4.3-GB Drives – RAID5
Network Interface	Compaq NetFlex-3 (100BaseTX)
Operating System	Windows NT Server v4.0 + SP2 Exchange Server 5.0

Table 3. ProLiant 6000 hardware configurations

PERFORMANCE RESULTS

Introduction

This section of the paper will focus on the performance of Microsoft Exchange Server in testing. Four main server resources will be addressed:

- Processor subsystem -- all the server's CPU resource, whether a uniprocessor system or a multiprocessor system
- Disk subsystem -- all the server's disk storage resource, including controller type and number of drives in a RAID set
- System Memory -- all the server's memory resource, but not including cache memory on processor boards or drive arrays. This is the amount of RAM installed in the system
- Network – the server's network connection and its capacity, performance and effect on Exchange Server response time

The cardinal rule to remember is this: ***If a subsystem is not a bottleneck, then adding more of that resource will not increase capacity***. In all likelihood, adding more of a non-bottlenecked resource will provide minimal improvements in response time, but not provide any additional system capacity.

When analyzing the different resource areas of a server (CPU, memory, disk, and network), it is also important to understand the performance-versus-cost trade-offs involved. If the addition of a resource provides an incremental ten-percent performance benefit, the cost must be weighed against the total system cost and the overall importance of the performance benefit to your organization. The important consideration is whether or not the performance gain justifies the additional cost.

Remember that costs can also come in the form of increased management and support overhead, system complexity, or more potential points of failure. An example of this with Microsoft Exchange is the question of adding a third processor. Exchange may benefit in response time to users by five to ten percent, but no additional system capacity (ability to handle more users) is achieved. In this case, it is unlikely that the additional cost for a processor is justified. Similarly, a large, high-end RISC-based system may provide performance gains of from five to ten percent, but may cost as much as 200% of the price of a Compaq ProLiant server. Again, the question is: Does the possible performance benefit justify the cost premium?

Response Time

Response time is the key measurement for servers that are directly supporting users in the Microsoft Exchange environment (for example, Mail and Public Folder servers). For servers that are supporting users indirectly (for example, Bridgehead, DL Expansion, Free/Busy, DS Replication), throughput is the more important indicator of performance. Servers indirectly supporting users present an additional level of complexity for capacity planning exercises and are outside the scope of this paper.

NOTE: Response time scores for Configuration A at 2500 and 3000 user loads were omitted since they failed to meet the sub-second (<1000 ms) requirement. In addition, Configuration B achieved a response time score of 1075 ms for 3200 users, which was just beyond the upper limit.

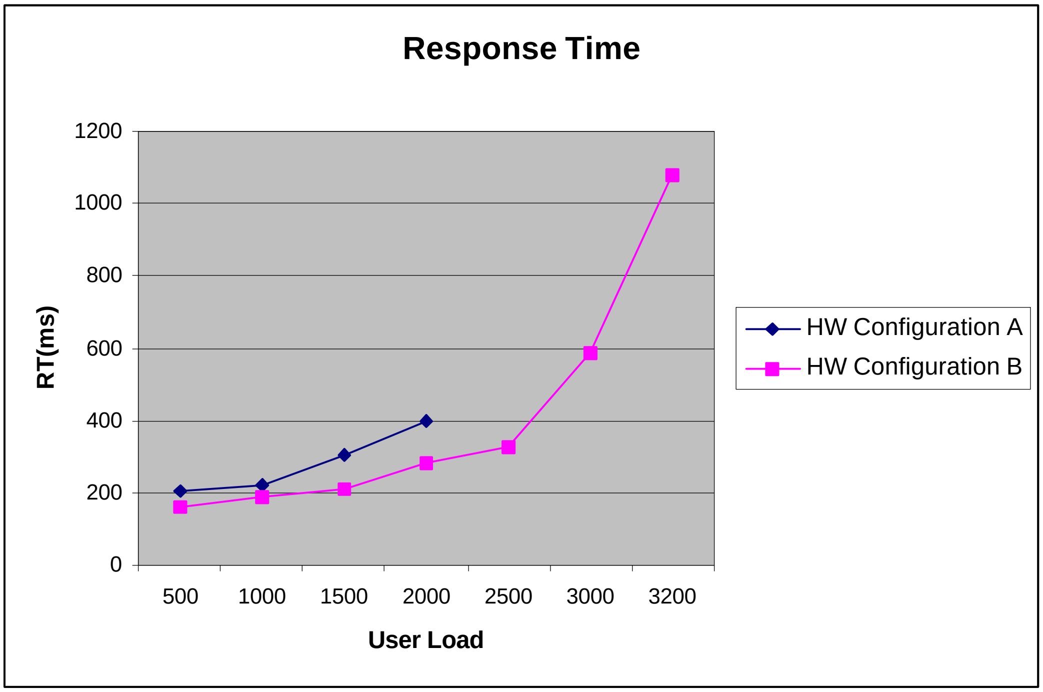

Figure 2. ProLiant 6000 Response Time (See Appendix for detailed information)

Processor Subsystem

In general, the processor subsystem is not an area of contention for Exchange servers directly supporting Exchange Mail users. However, certain specialized server functions in Exchange Server such as Public Folder, Bridgehead, DL Expansion, Free/Busy, and DS Replication do place heavier loads on the processor subsystem. Unfortunately, constraints imposed by the current version of Microsoft Exchange Server prevent its scalability in multi-processor systems beyond two CPUs. Here are some important points about processor scalability and utilization with Exchange Server.

- As mentioned, CPU scalability is limited to two processors. Additional processors may be added but will provide diminished returns (approximately ten percent or less). An additional CPU is recommended when there are more than 1000 users.
- Specialized Servers (see above) may require additional processor bandwidth.
- Limits imposed by Microsoft Exchange Server technology prevent further processor scalability at this time. NOTE: This issue will be addressed in the forthcoming version of Exchange.

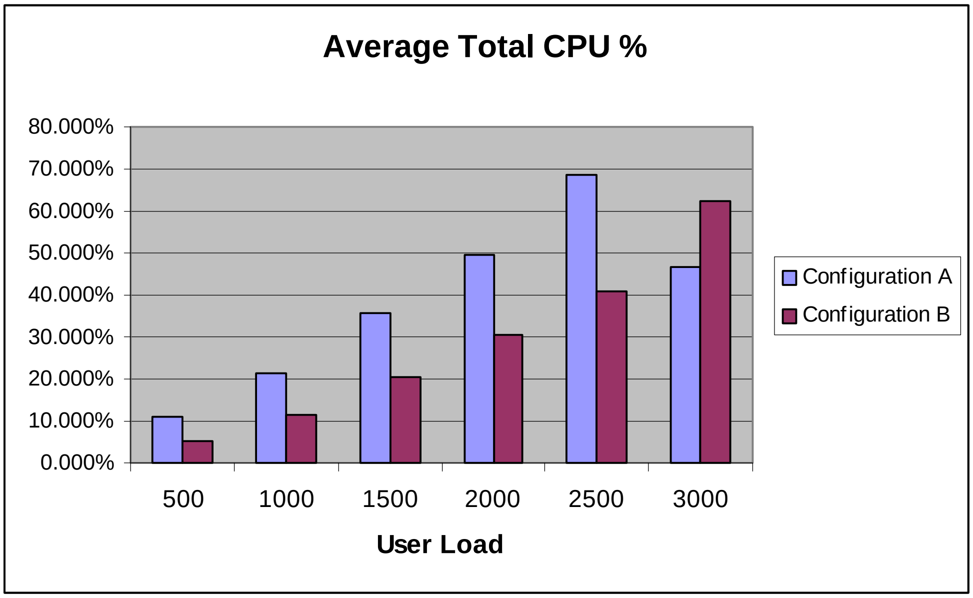

Figure 3. Total CPU Utilization at various user loads

NOTE: CPU scalability in Configuration A was limited by disk I/O at 3000 users. The processor was waiting on the disk subsystem (see Disk Subsystem). When the disk is no longer a bottleneck (Configuration B), CPU scaling is greatly improved (up to two processors).

A Note on Processor Technology

The Pentium Pro processor represents Intel's latest innovation and engineering in multiprocessor technology. With the advent of the Pentium Pro, X86 processor technology has been brought on a par with competing RISC technologies. Integer performance (what we care about in client/server applications) in the mainstream Pentium Pro is equivalent to or greater than that of competing technologies such as MIPS, PowerPC, and Alpha.

The Pentium Pro boasts such performance-enhancing technologies as Superscalar Pipelining, Branch Prediction, Out-of-order Execution, and Register Renaming that have made it the volume processor technology in the industry today. The X86 application base is one of many reasons that Compaq has built its entire server line around the Pentium and Pentium Pro processor technologies

A further important point is that a "chip" does NOT make a server. Contrary to what competing RISC-based server vendors would like the industry to believe, a server should be engineered for performance throughout all subsystems (not just the processor subsystem) and integrated with industry-leading applications such as Microsoft Exchange Server. These design goals have made Compaq the leader in the Intel-based server market.

Disk Subsystem

Effect of Number of Spindles

The basic idea with a disk subsystem is to provide the highest level of disk I/O possible to meet the application's requirements. One way to do this is to have many spindles (disks) connected to a Compaq SMART-2 Array Controller and employ multiple controllers. Each SMART-2 Controller can support up to fourteen drives across two Wide-Ultra SCSI channels. The SMART-2SL, however, supports a maximum of seven drives on a single Wide-Ultra SCSI channel. If each disk supports an average of 55 Random I/Os per second, a single SMART-2 Array Controller can deliver approximately 770 Random I/Os per second. For Sequential I/O, more than 2000 I/Os per second is possible (assuming approximately 150 sequential I/Os per second per drive).

Additional information helpful for capacity planning and sizing of client/server applications such as Microsoft Exchange Server is available in the following Compaq white papers:

• ***SMART-2 Array Controller Technology (Document number 317A/0797)***

• ***Configuring Compaq RAID Technology for Database Servers (Part number 184206-001)***

• ***Configuration and Tuning of Microsoft SQL Server 6.5 on Compaq Servers (Document number 415A/0696)***

The concepts discussed in these white papers are applicable to Microsoft Exchange Server. Microsoft Exchange is a high-transaction-throughput database engine based on Microsoft's *Joint Engine Technology* (JET). Many generic techniques and recommendations for tuning of database environments can be successfully applied to Microsoft Exchange Server. For example, the separation of sequential (for example, Log and Pagefile) I/O operations from random (for example, database access) I/O operations is an important optimization technique for both Microsoft SQL Server and Microsoft Exchange Server.

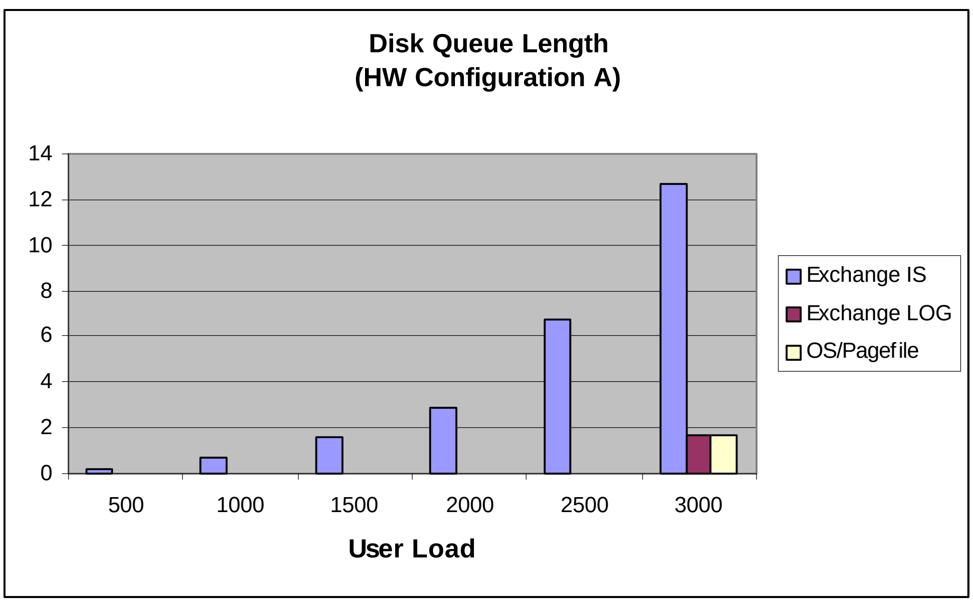

Figure 4. Disk Queue Length (Configuration A)

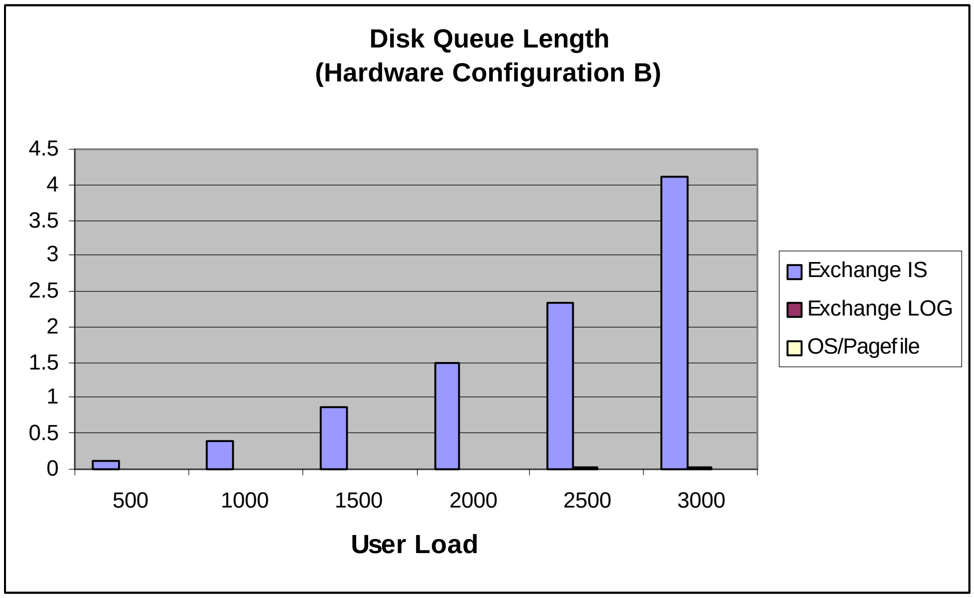

Figure 5. Exchange Disk Queue Length (Configuration B)

Disk I/O Profiling

An important technique in capacity and performance planning and analysis is I/O profiling. It is necessary to plan for the disk I/O capacity that your application will require in production. Once again, workload characterization is key to making this planning successful. With proper workload characterization and simulation, the I/O requirements for the disk subsystem can be observed, and the subsystem design can reflect those observations.

Microsoft Exchange Server is, in a generic sense, a client/server transaction-based database application. As such, the disk I/O profile is much like that of a database engine such as Microsoft SQL Server. Microsoft Exchange Server stores data in two key files, PRIV.EDB and PUB.EDB, which hold private data (Mail) and public data (Public Folders). A third file, DIR.EDB stores Exchange Server directory information. Two additional areas of interest are Queues (IMC, MTA, etc.) and Log files.

It is important to understand the types of disk I/O generated by access to each of these key data file areas in Microsoft Exchange Server. For the Exchange Server data files (PRIV, PUB, and DIR), the access patterns are random READ and WRITE (approximately 70% READ, 30% WRITE) in nature. For the Log files, the access patterns are sequential WRITE in nature.

This information is key to designing your disk subsystem to deliver the I/O capacity required by the production user load. For example, understanding not only the volume of disk I/Os but also the type (READ or WRITE) is a requirement for making choices about RAID levels and number of drives (See Compaq White Paper ***Configuring Compaq RAID Technology for Database Servers***).

NOTE: READ I/O on the IS for Configuration A decreased at 3000 users due mainly to disk bottlenecks and architecture issues in Exchange Server.

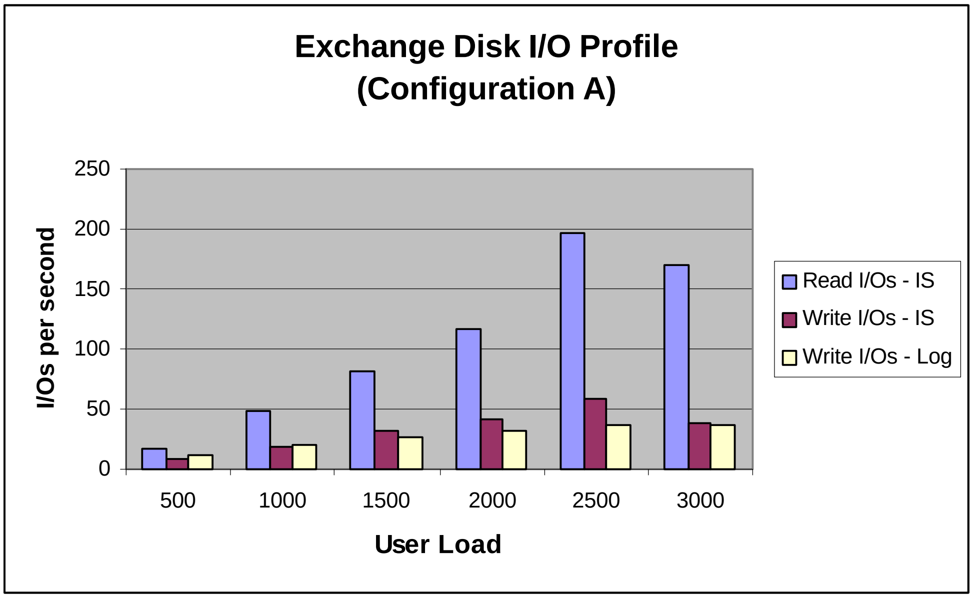

Figure 6. Exchange Disk I/O Profile (Configuration A)

NOTE: The Log files do not incur READ activity during normal operations, only during recovery.

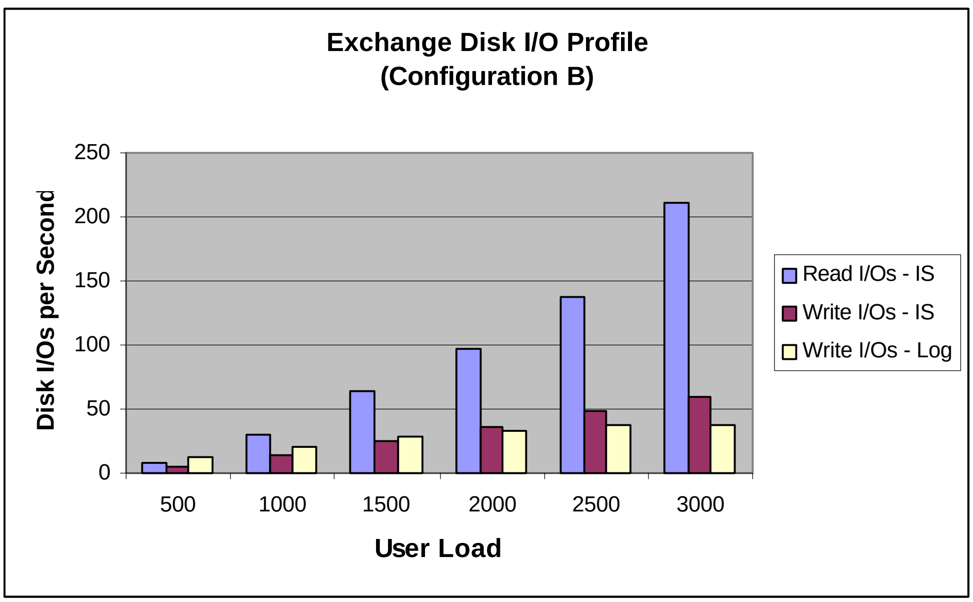

Figure 7. Exchange Disk I/O Profile (Configuration B)

***CONTROLLER CACHING**: Microsoft recommends turning off disk controller caching based on the assumption that common generic array controllers provide NO protection against data loss and corruption. The Compaq SMART-2 Array Controller, however, provides Error Checking and Correction (ECC) and battery backup for its cache. When a cache is amply protected, as in the case of the SMART-2, Compaq recommends enabling this cache for maximum performance benefit.*

Here are some important points about disk subsystem capacity planning and sizing:

- The benefit increases as the number of users increases. This is because the number of disk I/Os required also increases, so the benefit of the array controller becomes more apparent (compare Configuration A to Configuration B).
- Since the Exchange Server IS Buffers manage database/disk READ caching, Compaq recommends setting the SMART-2 Array Controller cache to 100% WRITE.
- The disk subsystem should be carefully designed in accordance with the I/O profile generated by the characterized workload.

- The disk subsystem becomes a bottleneck when the disk queue length exceeds from three to five I/Os outstanding.

NOTE: The SMART-2 cache can be tuned for WRITE or READ performance, although READ tuning will be beneficial only in applications that perform sequential READs.

Memory Subsystem

System Memory plays a crucial role in the performance of Windows NT and NT applications. It is very important to have enough system memory in an Exchange Server computer. For best performance, both Compaq and Microsoft recommend designing your system for NO operating system paging. In other words, the server should have enough physical memory to accommodate the application requirements in addition to enough free memory for Windows NT. Excessive paging results in drastic performance degradation due to unnecessary disk I/O activity.

The key memory resource area in Microsoft Exchange Server is IS Buffers. IS Buffers are the 4K pages that Exchange Server uses to cache information from the information stores. Rather than retrieving information from disk and incurring disk I/O, Exchange Server uses IS Buffers for better performance. Exchange Server does not use standard NT caching services.

During installation of Exchange Server, the Performance Optimizer is run to tune the Exchange Server based on the available hardware resources and performance characteristics. When tuning IS Buffers, PerfWiz (as it is called) can select a maximum of 20,000 IS Buffers by default. Therefore, Exchange Server can utilize a maximum of 80 MB of system memory for IS Buffers (20K * 4K per buffer). The 20,000-buffer limit is imposed by current Exchange architecture and will be addressed in the next release of Exchange Server.

Effect of Increasing IS Buffers

Aside from the amount of memory required for running Windows NT and Exchange Server, system memory and disk performance are interrelated. This is because the IS buffers are allocated from system memory. A small IS buffer pool will constrain disk I/O, whereas a large IS buffer pool will tend to reduce the disk I/O demands on the disk subsystem. The following points should be considered:

- Adding extra IS Buffers produces significant benefit, especially at higher user levels. This is because the extra IS buffers relieve the disk subsystem of a certain amount of I/O load.
- Most of the benefit from extra IS Buffers will be from READs. All WRITEs have to make it to the disk sooner or later, so the net number of WRITEs does not change greatly. However, significant reductions in READ I/O are observed.
- Depending on the cost of RAM versus the cost of extra disks, adding RAM may be a cost-effective alternative to adding disks to a RAID set in order to increase performance.

Effect of Increasing IS Buffers Too Much

It is possible to manually set the IS Buffers using the PerfWiz "–v" option. If the IS buffers are increased by too much without adding extra RAM, you will cause the system to become memory constrained and start paging. This will defeat the entire purpose because response time will degrade as a result.

NOTE: There has been no measured benefit to adding additional system RAM and increasing IS Buffers beyond 20,000. Although adding additional buffers has been seen to relieve disk I/O, current architecture limits in Exchange Server prevent taking advantage of these additional resources.

IS Buffers should be increased (up to a maximum of 20,000 buffers) based on free memory available in the system, not simply as a result of adding more RAM. Be careful not to take away memory from system processes. There should be at least 10 to 15 MB of free memory in the system at all times. It is always safe to run Exchange Optimizer to get a good recommendation for setting the number of IS Buffers based on system memory.

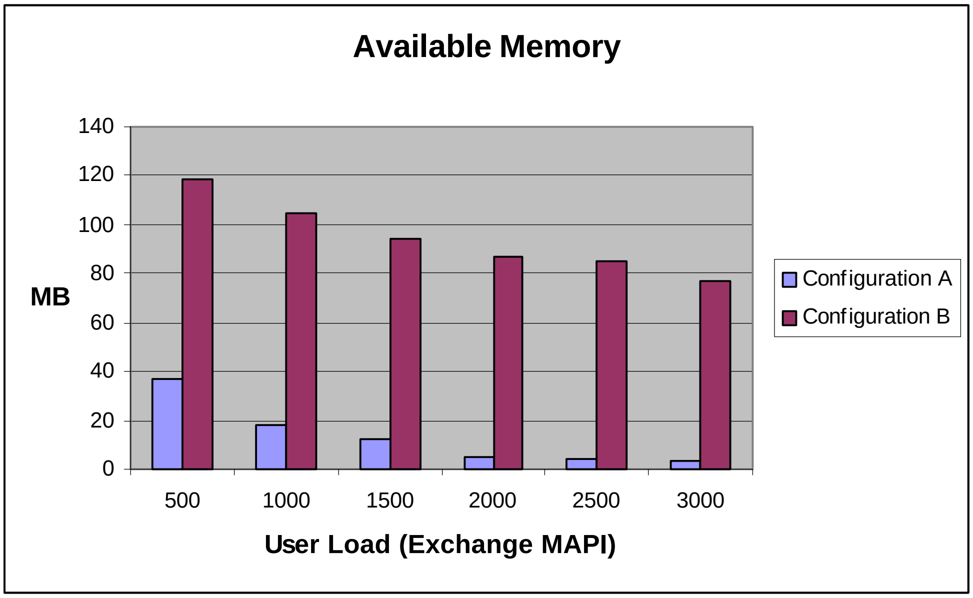

Figure 8. Available System Memory at various user loads

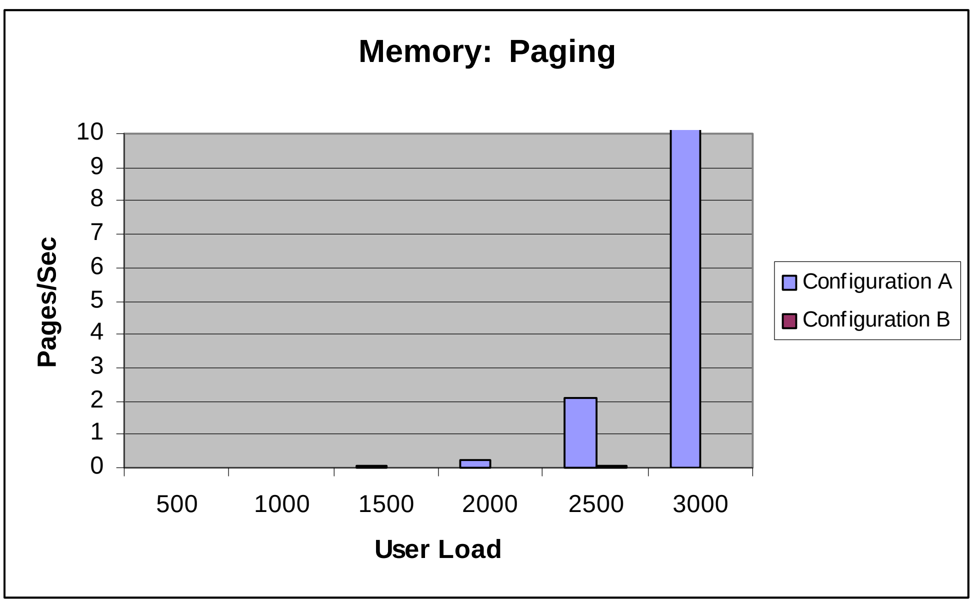

Figure 9. System Paging at various user loads and hardware configurations

Network Subsystem

NOTE: Testing by both Microsoft and Compaq has shown that client/server RPC traffic is not an area of great concern. In our simulations, even loads as high as 3000 users showed a network utilization of approximately 15% for 10Base-T (10 megabit) and approximately 2% for 100Base-TX (100 megabit).

The network subsystem in many respects is a non-issue in Exchange Server performance. Every indication is that, while network infrastructure should be an important consideration for Exchange Server planning and deployment, network bandwidth is not an area of resource contention with Exchange Server.

Exceptions to the above are in two areas. The first concerns the support of users over WANs and remote sites via slow links. Careful analysis and planning should be done when deploying Exchange Server to user populations over wide-area or slow links. Previous recommendations regarding proper workload characterization and analysis are applicable. The area covered by this exception is outside the scope of this paper, but several resources are available on the subject from Microsoft Product Support Services (PSS) and Microsoft Consulting Services (MCS).

The other exception involves Exchange Server remote procedure call (RPC) traffic. The fundamental communication mechanism in Exchange Server is the RPC. Both Exchange servers and Exchange clients use RPCs to communicate. Of concern for capacity planning purposes is inter-server RPC traffic. Exchange servers use RPCs to conduct such activities as DS replication, PF replication, and mail transfer. In the design of Exchange Server sites, careful analysis should be carried out to plan for RPC traffic between servers. For example, a dedicated backbone between servers is one option for ensuring adequate bandwidth for inter-server RPC traffic. Once again, Microsoft has dedicated substantial research and testing to these issues in order to assist customers in addressing them.

Conclusions and Recommendations

The purpose of this section is to provide general guidelines for configuring an Exchange Server computer. You should read it and apply the data in this paper to your own configuration decisions. It is organized into four categories: processor subsystem, disk subsystem, memory, and general.

Processor

- Typically, the processor subsystem is not the key resource bottleneck in Exchange Server 5.0. However, if the server will be supporting many users, and especially if the server will be handling many different non-IS processes such as Public Folders, Free/Busy, DL Expansion and DS Replication, multiple CPUs will be a benefit.
- The Pentium Pro has clear advantages over Pentium-based systems running Exchange Server. Pentium-based servers are not recommended for user loads greater than 1000.

Disk

- Place the database logs on a separate physical volume from that occupied by the Information Service databases. Using the two ports of a SMART-2 controller for this works well, as the SMART-2/P and SMART-2DH can perform simultaneous I/O on both ports.
- The database store (IS) volume should be composed of an array of at least three disks, and it should be fault tolerant. Due to the random nature of this volume's I/O, RAID 5 is a good fault tolerance configuration. Although RAID 5 imposes a performance penalty for disk WRITEs, the ratio of disk READs to disk WRITEs is about 3:1. This ratio can change if you have a very large IS buffer, but it generally holds true.

- The log volume must *always* be fault tolerant. Because of the sequential-WRITE-only nature of the log volume, having many disks in a RAID set does not provide a significant benefit. Two high-capacity disks mirrored (RAID 1) is a configuration that provides good sequential-WRITE performance and excellent fault tolerance. Note that the IS buffers will aid log volume performance.
- SMART-2 controllers with WRITE cache should be set to 100% WRITE (using the Array Configuration Utility) for all volumes. The SMART-2/P and the SMART-2DH have 4 MB and 16 MB of cache, respectively, that can be configured for WRITE. The SMART-2SL comes with 6 MB of cache that can be configured only for READ. The nature of Exchange Server disk I/O is sequential WRITEs for the Log files and random READs and WRITEs for the IS database files (READ-to-WRITE Ratio = approximately 3:1). Since the IS Buffers account for disk READ caching, and this I/O is random and small in size (4K), there is no benefit in setting the SMART-2 cache for READ performance.

Memory

NOTE: Remember that there is no benefit in tuning IS Buffers beyond 20,000 in Exchange Server 5.0.

- 128 MB of RAM is an adequate amount of system memory to start with, up to the 1000-to-1500-user range. 256 MB can provide additional benefit for user loads beyond 1500.
- Adding memory beyond 256 MB will provide minimal performance or capacity benefits.
- Monitor the available memory on the system using NT Performance Monitor. If there is unused memory available on a consistent basis, allocate a portion of it to the IS Buffers. However, do *not* over-allocate RAM. There should be at least 10-15 MB of free memory in the system at all times.
- Increasing RAM in the system and allocating it to the IS Buffers can significantly improve client response time. In some cases, doing this may be more cost-effective than adding disks to the drive subsystem.

General

- Always run Exchange Optimizer after the initial setup of Exchange Server. Also, run it after changing configuration of the server.
- Always ensure that your capacity planning simulations include proper workload characterization. Use the tools available to profile your actual messaging/GroupWare environment. These tools include STORSTAT, MAILSTORM, INETLOAD or other third-party tools.
- Use resources such as field staff, tools, and white papers available from both Microsoft and Compaq to assist with capacity planning, performance optimization, and deployment of Compaq servers in an Exchange Server environment.
- Be sure to consider whether a resource is a bottleneck before adding more of that resource.
- Understand what measurement is important for the appropriate capacity planning exercise. For example, response time is the key indicator of performance for servers directly supporting users (for example, Mail and Public Folder servers), whereas throughput is the key measurement for servers indirectly supporting users (for example, Bridgehead, DS Replication, Free/Busy, and DL expansion servers).

CONCLUSION

The Compaq ProLiant 6000-class server is an ideal deployment platform for Microsoft Exchange Server 5.0 in the enterprise. As this paper demonstrates, the ProLiant 6000 comfortably supports the maximum user loads attainable with the current version of Exchange Server. Future enhancements by Microsoft to Exchange Server will allow customers deploying Microsoft Exchange Server to take full advantage of Compaq's server architecture. Compaq's continued investment in optimization and integration of Exchange Server and other Microsoft BackOffice products on its entire server line will ensure successful deployment of Compaq servers by customers worldwide.

APPENDIX: FULL DISCLOSURE OF TEST RESULTS

RESPONSE TIMES

CONFIGURATION A	500	1000	1500	2000	2500	3000	3200
LoadSim Score (ms)	203	223	305	398	2051	8254	N/A
CONFIGURATION B	**500**	**1000**	**1500**	**2000**	**2500**	**3000**	**3200**
LoadSim Score (ms)	159	189	209	283	324	584	1079

CPU INFORMATION

CONFIGURATION A	500	1000	1500	2000	2500	3000
% CPU – STORE	7.42%	15.72%	27.85%	38.87%	54.82%	37.34%
% CPU – DS	0.08%	2.22%	2.05%	2.92%	3.62%	2.74%
% CPU – MTA	1.25%	1.41%	3.22%	4.37%	5.42%	2.02%
CPU Avg. Priv. Time	2.49%	4.16%	6.49%	4.16%	14.05%	13.18%
CPU Avg. Interrupt Time	0.13%	0.26%	0.42%	0.58%	0.86%	0.88%
Avg. Context Switches/sec	380	682	1040	1439	2334	2350
Avg. CPU Queue Length	2.34	2.505	2.859	3.495	4.879	4.051
CONFIGURATION B	**500**	**1000**	**1500**	**2000**	**2500**	**3000**
% CPU – STORE	6.6%	15.7%	31.4%	49.0%	67.5%	96.8%
% CPU – DS	0.7%	2.2%	2.3%	3.0%	3.6%	4.6%
% CPU – MTA	1.2%	1.4%	3.5%	4.6%	5.6%	7.4%
CPU Avg. Priv. Time	1.3%	2.8%	5.1%	7.9%	11.2%	19.2%
CPU Avg. Interrupt Time	0.1%	0.2%	0.3%	0.4%	0.6%	0.8%
Avg. Context Switches/sec	854	2168	3793	7143	10652	17957
Avg. CPU Queue Length	0.232	0.222	0.667	0.667	0.949	2.07

DISK INFORMATION

CONFIGURATION A	500	1000	1500	2000	2500	3000
Disk Queue Length – DB	0.201	0.704	1.581	2.91	6.76	12.718
Disk Queue Length – Log	0.005	0.009	0.012	0.016	0.034	1.703
Disk Queue Length – OS	0.005	0.009	0.012	0.016	0.034	1.703
Read I/Os – DB	16.36	48.55	80.87	116.22	196.85	169.29
Write I/Os – DB	8.53	18.4	31.015	42.05	58.43	37.92
Write I/Os – Log	12.33	20.66	26.774	32.14	36.62	36.43
CONFIGURATION B	**500**	**1000**	**1500**	**2000**	**2500**	**3000**
Disk Queue Length – DB	0.099	0.391	0.883	1.48	2.323	4.1
Disk Queue Length – Log	0.005	0.009	0.012	0.014	0.016	0.016
Disk Queue Length – OS	0.001	0.001	0.002	0.002	0.003	0.003
Read I/Os – DB	7.62	29.6	63.87	96.869	138.27	211.24
Write I/Os – DB	4.77	14.72	24.53	35.23	48.15	59.71
Write I/Os – Log	12.44	21.03	27.95	32.97	37.06	38.06

MEMORY INFORMATION

CONFIGURATION A	500	1000	1500	2000	2500	3000
Avg. Avail. Bytes (MB)	37	18	12	5	3.8	3.5
Avg. Pages/sec	0.016	0.014	0.08	0.228	2.07	103.27
CONFIGURATION B	**500**	**1000**	**1500**	**2000**	**2500**	**3000**
Avg. Avail. Bytes (MB)	119	105	94	87	85	77
Avg. Pages/sec	0.008	0.021	0.011	0.008	0.089	0.028

EXCHANGE SERVER INFORMATION

CONFIGURATION A	500	1000	1500	2000	2500	3000
IS Private Send Queue Avg.	0.03	0.273	1.253	1.162	953	7551
IS Public Send Queue Avg.	0	0	0	0	0	0
MTA Work Queue Avg.	0.071	0.111	0.394	0.556	1.495	1.626
CONFIGURATION B	**500**	**1000**	**1500**	**2000**	**2500**	**3000**
IS Private Send Queue Avg.	0.04	0.111	0.253	1.879	0.919	3.301
IS Public Send Queue Avg.	0	0	0	0	0	0
MTA Work Queue Avg.	0.04	0.071	0.253	0.444	0.505	1.01