



SUN AND EXCHANGE SOLUTION BUNDLES

*Guide to optimizing the Sun Fire™ X4100 M2 or X4200 M2 servers and
the Sun StorageTek™ 2530 Storage Array for Microsoft Exchange applications*

February 2008



Microsoft

Table of Contents

Optimizing Exchange with Sun	1
Business Drivers – The Balance of Strategy and Operations	1
The Sun Exchange Bundles	3
Sun Fire™ X4100 M2 and X4200 M2 – More than a Pizza Box	4
Sun SAS HBA for HA and non-HA Storage Connectivity	5
Sun StorageTek 2530 – Faster Exchange Queries	6
Steps for Optimizing Exchange	7
Define and Assess	8
Preparation and Installation	10
Sun Fire™ X4100 M2 and X4200 M2 Configuration for Exchange	11
StorageTek™ 2530 Configuration	12
Benchmark and Adjust	14
Appendix A – Exchange Benchmarks (Loadsim and Jetstress)	15
Exchange Benchmark Configuration	15

Optimizing Exchange with Sun

Microsoft Exchange® is the most widely used email application today, and expected to garner over 70 percent share by 2010, according to Gartner¹. To say that email has become mission critical is, to say the least, an understatement. Email is at the core of every business. Beyond the internal use of email, BlackBerry's and other remote email devices have helped corporate email reach across the globe for almost every company. Exchange is a database driven application and it has a specific set of CPU, memory and storage requirements to yield optimal performance.

The Exchange community is in the midst of a major transition from Exchange 2003 to Exchange 2007. Most IT managers are taking a measured approach to this transition. However, there is no question the transition is underway. Sun has teamed with AMD to build 64-bit, high-performance server/connectivity/storage solution bundles to meet the demands and needs of Exchange 2003 or 2007 for use in SMEs and departments. The word "optimized" means different things to different IT managers. In this paper, we provide framework for assessing, installing, and configuring Exchange environments to run faster, safer, and more efficiently on Sun's AMD-based platforms.

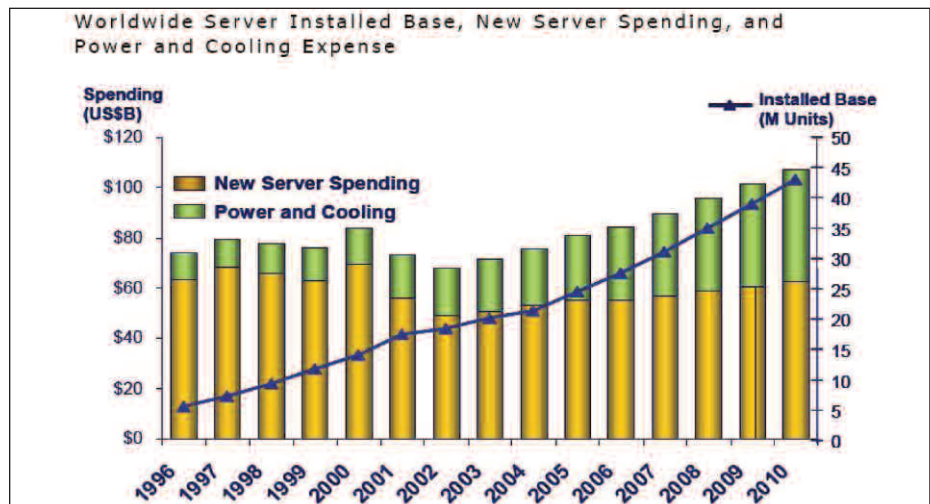
Business Drivers – The Balance of Strategy and Operations

More than ever, IT is the engine that drives business operations and email is one of the core elements of those operations. Email has become woven into the very fabric of our businesses. The role of IT often shifts between strategic and operational depending on the program in question. This year transitions to major Microsoft Platforms, Web 2.0, Green IT, Virtualization & Consolidation and providing better IT Service Level Agreements (SLAs) at lower costs are high on the list of topics.

- **New Microsoft platforms** – The next generation of Microsoft platforms (Windows 2008, Exchange 2007, and SQL 2007) need to be factored into today's purchasing decisions. Moving toward 64-bit architectures, choosing next generation storage models and server virtualization strategies are vital for IT managers.

¹ www.serverwatch.com/trends/article.php/3682041

- **Virtualization and consolidation** – Nothing has had a more dramatic effect on IT in the last two years than virtualization. Like a good smashup application that brings new value to old data, virtualization, power and cooling costs, and multi-core CPUs have created a “Perfect Storm” of opportunity. These technologies have appeared at just the right time, addressing the ongoing need to improve IT efficiency and support, processing more transactions, storing more data, and supporting more customers with lower overhead.
- **Green IT** – Green IT is not just about saving the planet, it’s about the real hard dollar costs of power, cooling, and rack space in datacenters. In many IT shops in major metropolitan areas, such as London or in the financial district of New York City, extra power is not available, or has multi-year waiting lists. According to IDC, power and cooling costs will exceed \$40B (USD) by 2010. IT managers need to partner with vendors who can help them build a Green IT strategy.



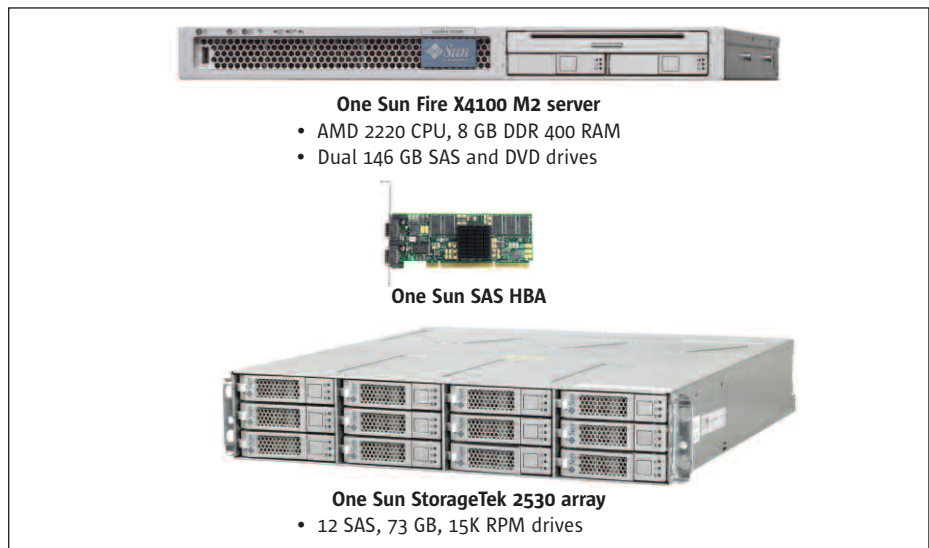
The growth of power costs in IT. Source: IDC, 2006.

- **Higher SLAs, for less money** – This is a theme that never goes away in any business. Global competition is relentless. If your company does not optimize, then a new company will. The secret to higher SLAs for less money is picking solutions that are optimized, designed for lights out operation and support by a global player, ready to deploy and ready to make the IT manager’s life easier.
- **Web 2.0** – Wiki’s, smashups, and social networking for business are at the forefront of the Web 2.0 movement. At the core of most of these applications lie fast efficient and flexible databases. These tools are growing because they help businesses collect and leverage the “tribal knowledge” to make the entire organization smarter and more effective. They help bridge groups of people to solve real business challenges for development and customer support.

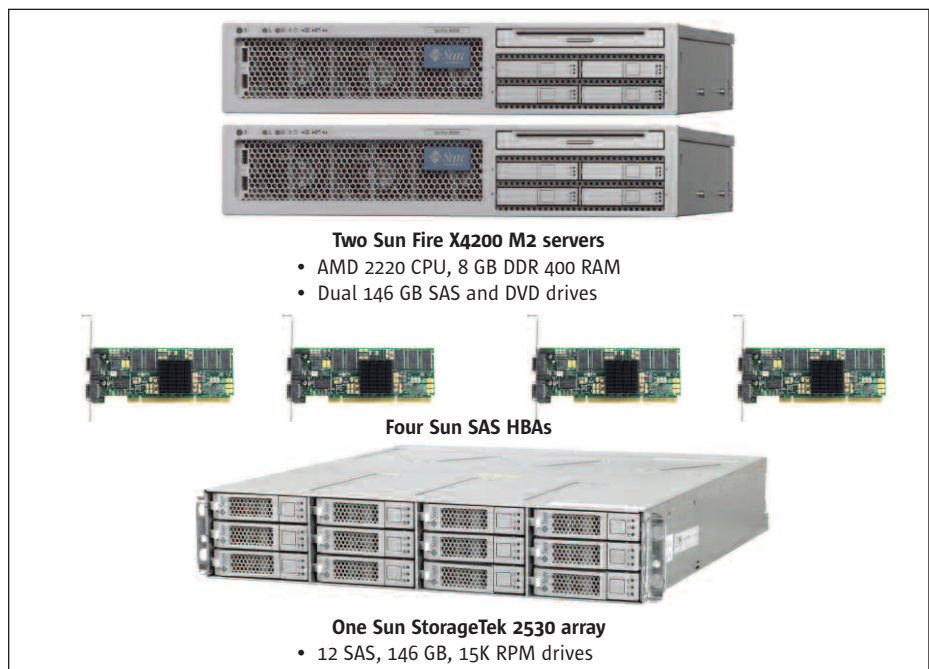
Sun is working with leading technology partners to help IT managers meet these strategic and tactical needs. Sun and AMD are leaders in virtualization, consolidation, Green IT and providing higher levels of SLAs. The Exchange bundle is designed to simplify deployment and optimization to meet these growing IT needs.

The Sun Exchange Bundles

Sun and AMD are leaders in innovative performance computing. Together they have teamed to bring the world-class performance of the Sun Fire X4100 M2 and X4200 M2 servers and Sun StorageTek 2530 storage array in a pair of solution bundles that are ideal for SMEs and departmental needs. For added value, Sun recommends customers purchase a Sun Spectrum Service plan and take advantage of features such as hardware service coverage, telephone and online technical support, SunVIP interop support, and firmware updates to help reduce storage system management costs and improve performance. Bundle details for optimized non-HA and HA Exchange applications:



Exchange Bundle #1



Exchange Bundle #2

Sun Fire X4100 M2 and X4200 M2 – More than a Pizza Box

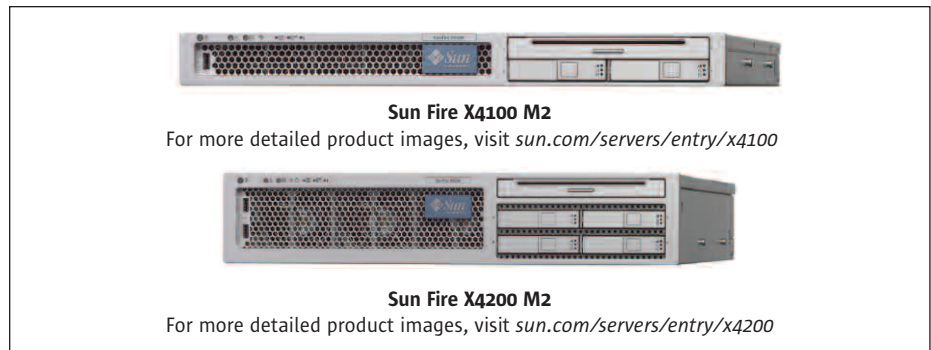
The venerable 1U server is often called the pizza box. Those are not words that describe the Sun Fire X4100 M2 server, a traditional 1U form factor (Brooklyn Style) and the Sun Fire X4200 M2 server, a 2U (Chicago Deep Dish) server. In selecting the solution bundle components from Sun's portfolio of products, we chose a best-in-class platform that helped set a world record for database performance. How do we define best-in-class for running Exchange?

- **More Exchange performance with HyperTransport memory** – For database applications, the faster the memory, the better it can perform transactions. The HyperTransport architecture of the X4100 M2 and X4200 M2 run Exchange faster with aggregate bandwidth up to 24 GB/s per processor. The AMD Opteron processor with HyperTransport provides a scalable direct connection between processors, I/O subsystems, and other chipsets. HyperTransport technology helps reduce the number of buses in a system, which lowers transaction overhead to improve Exchange transactions and queries.
- **Industry leading CPU performance²** – The AMD multi-core CPUs of the Sun Fire X4100 M2 and X4200 M2 servers deliver industry leading x86 CPU horse power. This enables greater consolidation of servers and more VMs of a single system, higher performance at a lower cost with managers improve SLAs, lower acquisition costs, and increased consolidation.
- **Low power and cooling demands³** – The AMD family of CPUs are more power efficient CPUs. The Sun Fire X4100 M2 and X4200 M2 servers also use a 2.5 inch Serial Attached SCSI (SAS) enterprise class disk drive that offers IOPS performance increase and 58 percent power savings compared to today's 3.5 inch drive. This helps IT managers build a Green IT strategy.
- **More connectivity for better virtualization** – The Sun Fire X4100 M2 and X4200 M2 servers are the only systems in this class with 4 native 1 GbE ports and embedded SAS ports on the base board. The extra connectivity on the Sun Fire X4100 M2 and X4200 M2 servers increase systems accessibility by clients and make it more capable of hosting high performance applications, and supporting more VMs for a larger set of clients. This increased connectivity allows for more bandwidth, without using expansion slots and use less power.

² [sun.com/servers/entry/X4100 M2 and X4200 M2/](http://sun.com/servers/entry/X4100_M2_and_X4200_M2/)

³ [sun.com/servers/entry/X4100 M2 and X4200 M2/](http://sun.com/servers/entry/X4100_M2_and_X4200_M2/)

- **More flexible RAS and HA options** – The Sun FireX4100 M2 and X4200 M2 servers provide the most flexible hot-swap components in its class. This increased functionality, combined with our stand-alone ILOM, for lights-out operation makes the Sun Fire X4100 M2 and X4200 M2 servers the leader for HA and RAS in the x86 market.



Sun Fire X4100 M2 and Sun Fire X4200 M2 servers

Sun SAS HBA for HA and non-HA Storage Connectivity

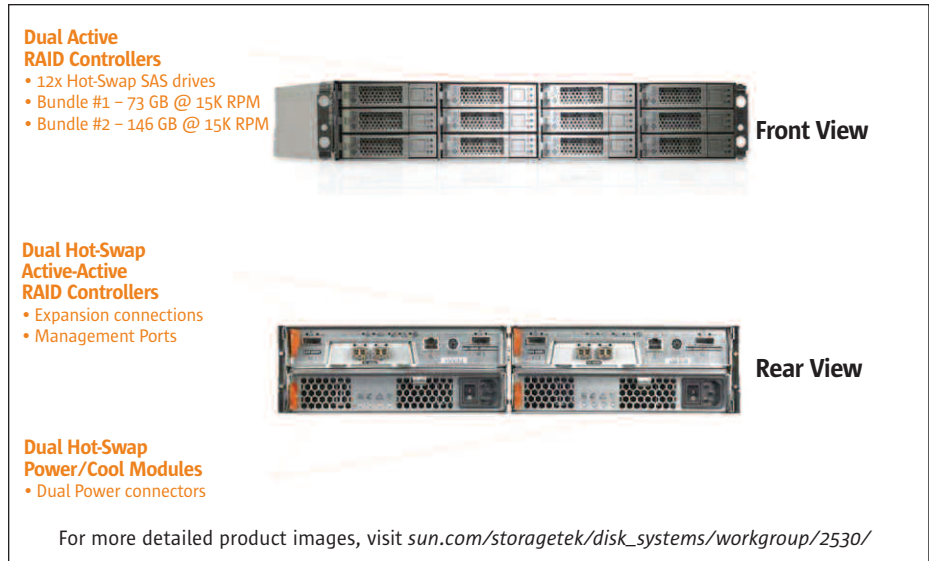
The Exchange HA bundle includes four high performance PCIe SAS Host Bus Adapters (HBAs) and cabling that provide redundant storage connectivity across both servers. This enables better load balancing and fail over using Sun tools or native OS features. The non-HA Exchange bundle includes a single high performance PCIe SAS HBA and cabling for easy connectivity to the Sun StorageTek 2530. Sun's PCIe SAS HBA addresses the need for SAS connectivity in the Exchange Bundle. This HBA offers outstanding performance by matching 8-lanes of PCIe, with 8-lanes of 3 Gb/s SAS. The SAS interface allows for point-to-point connection enabling higher performance, increased aggregate bandwidth, higher availability with dual-ported drives, and enhanced reliability when compared to parallel SCSI. The SAS interface provides universal interconnect with Serial ATA (SATA), while still offering logical SCSI compatibility along with the reliability, performance and manageability available in SAS drives. Common electrical and physical connection allows for the application to dictate the technology of choice for disk drives. The Sun StorageTek™ PCIe SAS Host Bus Adapter, 8-port, supports the following:

- Two x4 external miniSAS connectors (SFF-8088)
- Connects to both SAS and SATA HDDs
- Allows up to 244 end-point devices
- Enterprise-class HBAs ideal for large or mixed OS environments
- Features the award-winning Fusion-MPT HBA architecture from LSI
- Includes an installed low-profile bracket and full height bracket
- Simultaneous on-board support for x64/x86 BootBIOS (Windows and Linux) and Solaris™ SPARC OpenBoot

Sun StorageTek 2530 – Faster Exchange Queries

Storage is a key component of optimizing Exchange performance. For this bundle, we selected the SAS-based StorageTek 2530. SAS is the latest storage technology to be globally adopted for the enterprise, because it brings the high performance and reliability at a lower cost. The StorageTek 2530 is what SMEs and departments need to build mission-critical applications on a frugal budget. The Sun StorageTek 2530 array is an easy-to-use array that delivers better price/performance and greater rack density than comparable solutions. The StorageTek 2530 is a leading example of Sun's storage expertise to provide best-of-breed technology, reliability, and affordability.

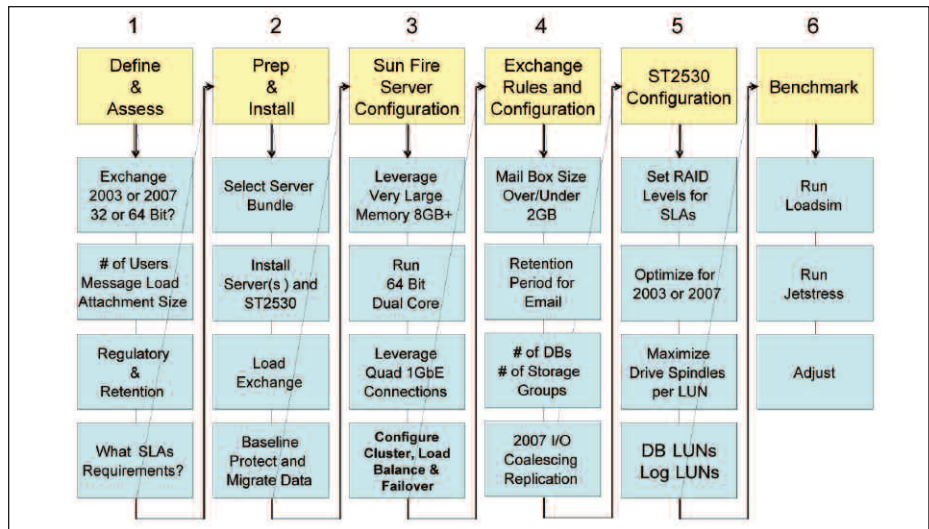
- **Easy to configure and manage for Exchange** – The Sun StorageTek 2530 is ready to support Exchange out of the box, making storage management as easy as server provisioning. The SMI-S-compliant Sun StorageTek Common Array Manager (CAM) software ensures a friendly user interface from setup to administration. The Web-based task management interface puts you in control of the array and significantly reduces the complexity of installation, configuration, and management. Online capacity expansion, data volume control creation, and host-to-volume mapping are just a click away, making it easy to make changes quickly and efficiently. CAM is accessible from anywhere in the world and includes automated diagnostics so you can focus on precise, predictable, repeatable results. CAM also scales across Sun's entire modular disk portfolio.
- **More Exchange transactions with 15K SAS Drives** – The StorageTek 2530 supports multiple sizes and speeds of drive. Exchange has multiple performance needs. The Jet and Streaming databases that underlie the application need storage that is optimized for random access, while log files require storage that is optimized for sequential transactions. The StorageTek 2530, bundled with 73 GB or 146 GB SAS drives, provide high performance write-through capability. Additional write-through performance can be added by expanding to 36 drives.
- **Data protection for Exchange** – Managing growth and expansion doesn't come easily, especially when you're operating within a budget. With dual-active RAID controllers, high availability, and high density, the Sun StorageTek 2530 array provides a cost-effective solution that lets you take advantage of improved performance and low-cost growth. The Active-Active configuration supports load balancing and failover for improved performance and HA.



Sun StorageTek 2530

Steps for Optimizing Exchange

Optimizing your Exchange environment takes planning and understanding. The number of users, the work load of the messages, the size and frequency of attachments, and the ratio of read vs. writes and local vs. BlackBerry users. Additionally, regulatory retention policies and SLAs effect configuration requirements. In this section we outline a way to evaluate the right settings for your application.



Steps for optimizing Exchange.

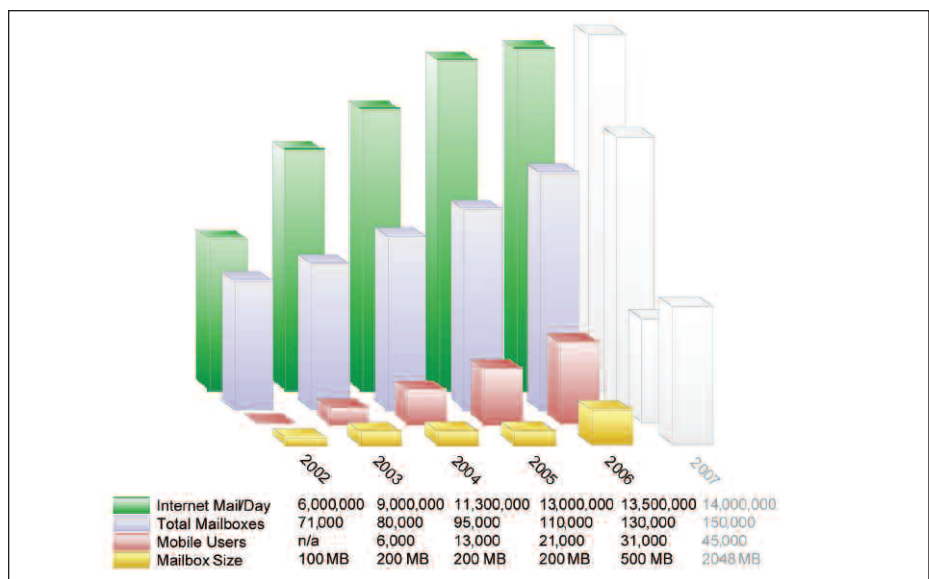
Define and Assess

Understanding the data is the first step in optimizing the bundled infrastructure to support Exchange. The nature of the read/write activity and ratios is vital to establish which settings will help optimize your solutions. The reason for this is to find the optimal buffers and caches to reduce excess memory and I/O traffic that are tied to the data characteristics.

- **Exchange 2003 or 2007** – The version of Exchange you choose has a fundamental impact on the configurations of both the server and storage. The following table outlines some of the changes that affect the optimization settings of both the server and the storage.

Function	Exchange 2003	Exchange 2007
Max memory	4 GB	16 GB/CPU
Mode	32-bit	32- or 64-bit
CPU	Multi-core	Multi-core
Max mail box	<2 GB	2 GB
I/O coalescing	NO	Yes
Page size	4K	8K
Log files	Over 1 MB	1 MB
Databases	20	50
Storage groups	20	50
Continuous replication	–	Local and remote

- **Number of users, message load and attachment** – Knowing this up front enables better understanding of the overall work load on the server and storage. The number of users and kind of users helps determine the message load. The graph below, from Microsoft, shows the growth of email message and mailboxes⁴.



Growth of email message and mailboxes.

⁴ Enterprise Messaging with Exchange 2007 Technical Brief Nov. 2006

Local users generate less traffic and a more consistent read/write ratio. The following table lists mailbox profiles that can be used as a guideline for capacity planning of Exchange mailbox servers. These profiles represent mailbox access for the “average user” Outlook (or MAPI-based) client within the organization.

User Type	Database Volume IOPS	Send/Receive per day	Mailbox size
Light	.5	20 sent/50 received	50 MB
Average	.75	30 sent/75 received	100 MB
Heavy	1.0	40 sent/100 received	200 MB
Large	1.5	60 sent/150 received	500 MB

Each profile represents total I/O to the Jet database and does not include I/O related to transaction log file activity. In order to accurately calculate your disk subsystem load, you must split this database I/O into read and write I/O because write operations are more I/O intensive than reads. To help estimate your own read/write ratio, consider the usage patterns of a company that has a heavy mailbox profile. In a production environment, the company can expect to incur read/write ratios between 75/25 percent and 66/33 percent, depending on the group of users being evaluated.

- For a mail system consisting of 2,000 heavily used mailboxes, a total of 1,500 IOPS is generated on the database volume. The formula to calculate this is:
 - In this example, .75 IOPS _ 2,000 mailboxes = 1,500 IOPS

Using a conservative ratio of two reads for every write (66 percent reads to 33 percent writes), you would plan for 1,000 read I/O and 500 write I/O requests per second for your database volume. Every write request is first written to the transaction log file and then written to the database. Approximately 10 percent of the total 1,500 IOPS seen on the database volume will be seen on the transaction log volume (10 percent of 1,500 is 150 IOPS); 500 write I/O requests will be written to the database.

Message load is with a significant number of remote emails devices such as BlackBerry’s cause more traffic per user. BlackBerry users cause additional overhead that affect the database IOPS of a server. When RIM tested 1000 BlackBerry enabled MMB2 users with BlackBerry Enterprise Server 4, they saw database IOPS increase by a factor of 3.64 over the standard MMB2 user without BlackBerry. This factor could be significantly smaller or larger depending on how BlackBerry devices are used in the environment. The BlackBerry test included: 10 synchronization commands; two memo adds, one modify, one delete; and four task adds. Actual BlackBerry device use will not be this constant, causing a lesser or greater affect on actual IOPS. For a mail system consisting of 2,000 heavily used mailboxes, of which 500 are BlackBerry enabled, a total of 3820 IOPS is projected on the database volume.

The formula to calculate this is:

Estimated BlackBerry IOPS per User for User Type _ Number of Users in this example, 1.0 IOPS _ 2,000 mailboxes=2,000 IOPS. If 500 of those users have BlackBerry devices, then those 500 users add 500 mailboxes x 3.64 IOPS=1820 IOPS, or 3820 total IOPS.⁵

- **Regulatory and retention requirements** – Is this Exchange server subject to any specific regulatory or security/access control requirements? Does the data need to be stored in a dedicated LUN? What are the business continuance requirements? Does the data need to be replicated synchronously or asynchronously, locally or remotely? What is your retention policy? It is important to understand these requirements before the optimization is planned.
- **SLA requirements** – As with regulatory requirements, SLAs drive the type of storage and data protection model that needs to be implemented on the array, how the data is backed up, how often the backup/snapshot is run. SLAs also help set the business continuance requirements for the application.

Preparation and Installation

The preparation for installing the new Sun Fire X4100 M2 or X4200 M2, and the StorageTek 2530 starts with four basic items:

- **Select server bundles** – For Exchange applications that are under 2000 users and do not require HA configurations we recommend the Sun Exchange Bundle #1 with one Sun Fire X4100 M2 server and the StorageTek 2530 loaded with 73 GB drives. Select Bundle #2 with two Sun Fire X4200 M2 servers if you require support for Microsoft Cluster Server or the additional capacity of 146 GB drives in the StorageTek 2530.
- **Physical location** – Finding the right rack space, power and cooling to support the hardware, and required switches and other infrastructure. This should be a simple, but required step.
- **Installing the Sun Fire X4100 M2 or X4200 M2, StorageTek 2530 and OS** – Sun provides a detailed step-by-step installation and configurations guide for each product at <http://docs.sun.com/app/docs/prod/5b03d0ed-216d-11db-a023-080020a9ed93>. Next is to load the OS of choice for the applications and set up the basic RAID and LUN configuration to hold new or migrated data.
- **Load Exchange** – The next step in the optimization process is to load Exchange onto the new Sun Fire X4100 M2 or X4200 M2 server and test access to the data on the new array.

⁵ Microsoft's "Optimizing Storage for Microsoft Exchange Server 2003" whitepaper, Dec. 12, 2006.

- **Baseline, protect, migrate and validate the data** – If this hardware is replacing a current server, make sure you have a complete backup of the data set to be moved, preferably to another disk array for rapid access. Next, migrate the data from the old server storage to the new StorageTek 2530 array and server. Finally, validate that the data set was moved properly, especially if you have to erase the backup and original copies.

Sun Fire X4100 M2 AND X4200 M2 Configuration for Exchange

The Sun Fire X4100 M2 and X4200 M2 servers are a true 64-bit system, and with AMD's HyperTransport memory performance, it makes Exchange run faster, hit more requests from systems RAM and reduce disk I/O. For more details on configuring the servers and memory for Exchange go to: <http://www.microsoft.com/technet/prodtechnol/exchange/2007/plan/hardware.mspx?wt.svl=highlights>

- **Leverage very large memory** – The Sun Fire X4100 M2 and X4200 M2 servers support up to 16 GB of RAM per CPU or up to 64 GB in a single server. The use of large memory reduce the number of disk I/Os the server needs to perform.
- **Run 64-bit dual core CPUs** – The uses of dual core CPUs can improve performance by 76 percent over single core CPUs. It also reduces the need for context switching that increase performance by reduce latency.⁶
- **Leverage quad 1 GbE connections** – The direct point-to-point HyperTransport architecture allows for higher I/O rates and helps the Sun Fire X4100 M2 and X4200 M2 servers support more clients in the same system.
- **Configure Microsoft Cluster, load balancing and fail over** – The Sun Exchange Bundle #2 includes two Sun Fire X4200 M2 servers. This allows for configuration of Microsoft Cluster server. As we have discussed, Exchange is a mission-critical application for every business and the use of clustering provides high-availability, scalability, load balancing of servers' workload and storage workload.

⁶ <http://www.microsoft.com/technet/prodtechnol/exchange/2007/plan/hardware.mspx?wt.svl=highlights>

StorageTek 2530 Configuration

One pundit said it best—“If the storage is down, Exchange is down...if the storage is slow “Exchange is slow.”

- **Configure RAID levels**⁷ – All Exchange data is stored in the Exchange store, which is composed of three major components. The following table lists the three major components of the Exchange store and their affect on disk I/O.

Component	Why it effects disk I/O
Jet database (.edb file)	<ul style="list-style-type: none"> • The Jet database is used to store all data submitted from MAPI clients. • All client activity generated by a MAPI client causes updates to the Jet database. • Stores incoming SMTP mail that contains MAPI information.
Streaming database (.stm file)	<ul style="list-style-type: none"> • Stores attachments and data submitted from IMAP4, NNTP, Microsoft Outlook® Web Access, or SMTP. Pointers are saved in the Jet database so the data can be delivered to MAPI clients upon request. • Stores incoming SMTP mail that does not contain MAPI information. • All Internet protocol client activity causes updates to the streaming database.
Transaction log files (.log files)	<ul style="list-style-type: none"> • All changes made to the database are first committed to transaction log files. This means that any time a user sends or reads a message, and any time a user modifies data stored in their mailbox, that change is written to the transaction log file. The change is immediately committed to the in-RAM database cache, and then copied back to disk when the system’s load permits. Transactions are also read back when a database is mounted.

Because each Exchange store component is written to differently, you will experience better performance if you place the .edb files and corresponding .stm files for one storage group on one volume, and place your transaction log files on a separate volume. The following table lists how the disk read/writes are performed for each Exchange store component.

⁷ Microsoft’s “Optimizing Storage for Microsoft Exchange Server 2003” whitepaper, Dec. 12, 2006

Component	I/O pattern
Jet database (.edb file)	<ul style="list-style-type: none"> • Microsoft recommends RAID 10 provides the optimal performance and protection by reduces write traffic (only two writes primary and secondary). • Read from and write to at random • 4 KB page size
Streaming database (.stm file)	<ul style="list-style-type: none"> • Microsoft recommends RAID 10 provides the optimal performance and protection by reduces write traffic (only two writes primary and secondary). • Normally read from and write to sequentially • Variable page size that averages 8 KB in production • Note: There are significant numbers of seek operations, so the I/O pattern is neither entirely random nor entirely sequential.
Transaction log files (.log files)	<ul style="list-style-type: none"> • Microsoft recommends RAID 10 provides the optimal performance and protection by reduces write traffic (only two writes primary and secondary) • 100 percent sequential writes during normal operations • 100 percent sequential reads during recovery operations • Writes vary in size from 512 bytes to the log buffer size

- **Sun RAID recommendation** – Sun has run a series of benchmarks and feels that RAID 5 provides a better balance of cost/performance. RAID 10 required the RAID system allocate twice the data capacity, where RAID 5 overhead ranges from 12-30 percent, depending on the number of drives in the stripe.

The table below shows, in the Exchange context, the RAID trade-offs as well as quota sizes available with configuration of both log and database volumes on this array.

- The first section shows the available quota if the entire tray was used exclusively for the Exchange database running RAID 10.
- The second section shows the available quota if the entire tray was used exclusively for the Exchange database running RAID 5.
- The third section which is more likely shows a RAID 10 implementation with both the log volume and database volume on the same array.
- The fourth section shows a RAID 5 implementation with both the log volume and database volume on the same array.

users	disks	size	raw disk	Only Exchange db		Only Exchange db		4 disk log 8 disk db Quota			3+1 Log 7+1 db Quota		
				RAID 10	Quota	RAID 5	Quota	RAID 10	RAID 10	RAID 10	RAID 5	RAID 5	RAID 5
2000	12	73000	876000	438000	219	727080	363.54	14450	293460	146.73	219000	511000	255.5
2000	12	146000	1752000	876000	438	1454160	7272.08	289080	586920	293.46	438000	1022000	511

- **Optimize for Exchange 2003 or 2007⁸** – Exchange 2003 has several well-documented performance challenges, (4K pages, individual page reads, memory limitations etc.). Exchange Server 2007 had added a number of items that make Exchange 2007 easier to optimize including:
 - I/O coalescing to reduce I/O demand by 30 percent vs. the same read/writes on 2003.
 - The Log files size reduction to 1 Mb makes replications easier and faster.
 - Continuous replications add data protection locally and remotely, but can double storage capacity requirements
 - The increase in data bases and storage groups make it easier to backup and group storage needs. It can also help support regulatory needs for email data bases the need to be separate. Note that the increase in databases and storage groups increase the amount of physical memory required. See this link for more details: microsoft.com/technet/prodtechnol/exchange/2007/plan/hardware.mspx?wt.svl=highlights
- **Maximize drive spindles and Split DB and Log LUNs with 15K SAS drives** – The StorageTek 2530 supports multiple sizes and speeds of drive. Exchange, like most database applications has multiple performance needs. The Jet and Streaming data bases that underlie the application need storage that is optimized for random access, the Log files require storage that is optimized for sequential transactions. The StorageTek 2530 is bundled with the 73 GB or 146 GB SAS drives that provide lower costs than Fibre channel and focus on increased write-through performance. Additional write-through performance can be added by expanding to 36 drives.

Benchmark and Adjust

The final step is to put it all together and make any final adjustments. Using tools like Loadsim and Jetstress to benchmark the performance gains to measure your improvements. Appendix A shows the internal Sun testing with the Sun Exchange bundle.

⁸ Microsoft's "Optimizing Storage for Microsoft Exchange Server 2003" whitepaper, Dec. 12, 2006

Appendix A – Exchange Benchmarks (Loadsim & Jetstress)

Exchange Benchmark Configuration

This document provides information on Sun Microsystem's storage solution for Microsoft Exchange Server, based the Microsoft Exchange Solution Reviewed Program (ESRP)-Storage program*. For any questions or comments regarding the contents of this document, see Contact for Additional Information.

*The ESRP-Storage program was developed by Microsoft Corporation to provide a common storage testing framework for vendors to provide information on its storage solutions for Microsoft Exchange Server software. For more details on the Microsoft ESRP-Storage program, please click microsoft.com/technet/prodtechnol/exchange/2007/esrp.msp

Disclaimer

This document has been produced independently of Microsoft Corporation. Microsoft Corporation expressly disclaims responsibility for, and makes no warranty, express or implied, with respect to, the accuracy of the contents of this document. The information contained in this document represents the current view of Sun Microsystems, Inc. on the issues discussed as of the date of publication. Due to changing market conditions, it should not be interpreted to be a commitment on the part of Sun Microsystems, Inc., and Sun Microsystems, Inc. cannot guarantee the accuracy of any information presented after the date of publication.

Features

This document describes a Microsoft Exchange primary storage solution on the Sun Microsystems StorageTek 2530 array. The targeted number of users in this environment is 4000. The tested user profile was 0.5 IOPS (0.42 + 20 percent) per mailbox with a limit of 300 MB. This solution can be scaled by adding additional arrays. The StorageTek 2530 is an excellent choice for medium- to large-sized businesses.

Solution Description

The tested solution consists of (1) Sun StorageTek 2530 RAID Unit with (12) 146 GB 15,000 rpm SAS 4 Gb/s drives and (1) JBOD Unit with (12) 146 GB 15,000 rpm SAS 4 Gb/s drives for 3.5 TB (raw)/1.8 TB (usable, RAID 10). This solution can expand to (2) JBOD Expansion units for a total of 5.26 (raw)/2.63 TB (usable, RAID 5 with hot spares) Fibre Channel or 420 TB (raw)/292 TB (usable, RAID 5 with hot spare) SATA-II. For more detailed specs, please see: sun.com/storagetek/disk_systems/workgroup/2530/

The Sun StorageTek 2530 Array is listed on the Windows Server Catalog of tested products here:

windowsservercatalog.com/item.aspx?idItem=53b769e5-acbd-4119-7d29-27a55a49306f

RAID Module LUNs

Sun StorageTek 2530 RAID Module LUNs can be managed through Sun StorageTek Common Array Manager (CAM). RAID levels 0, 1, (1+0), 3 and 5 are supported.

Tested Configuration

The LUN layout of this solution is as follows: Two LUNs per tray were created; one LUN was created for sequential writes (4 drive RAID 10) and a second volume (8 drive RAID 10) for random access. This solution is comprised of two 4-drive RAID 10 volumes for the logs and two 8-drive RAID 10 volumes for the database. The I/O activity to the log is 100 percent small sequential writes. The I/O activity to the database is mostly random. By separating the two distinct I/O characteristics to different LUNs, performance will increase.

The ESRP-Storage program focuses on storage solution testing to address performance and reliability issues with storage design. However, storage is not the only factor to take into consideration when designing a scale-up Exchange solution. Other factors which affect the server scalability are: server processor utilization; server physical and virtual memory limitations; resource requirements for other applications; directory and network service latencies; network infrastructure limitations; replication and recovery requirements; and client usage profiles. All these factors are beyond the scope for ESRP-Storage. Therefore, the number of mailboxes hosted per server as part of the tested configuration may not necessarily be viable for some customer deployment.

For more information on identifying and addressing performance bottlenecks in an Exchange system, please refer to Microsoft's Troubleshooting Microsoft Exchange Server Performance, available at <http://go.microsoft.com/fwlink/?LinkId=23454>

Targeted Customer Profile

This solution is targeted towards a 4000 user Exchange 2007 deployment.

- Medium to large organization
- One Exchange 2007 Server
- .42 IOPS + 20 percent headroom User I/O profile (0.5 tested)
- 300 GB mailbox quota

Tested Deployment

The following tables summarize the testing environment:

Simulated Exchange Configuration

Number of Exchange mailboxes simulated	4000
Number of hosts	1
Number of mailboxes/host	4000
Number of storage groups/host	2
Number of mailbox stores/storage group	1
Number of mailboxes/mailbox store	2000
Number of mailbox store LUNs/storage group	1
Simulated profile; I/Os per second per mailbox (IOPS, include 20 percent headroom)	.5 (very heavy workload)
Database LUN size	540 GB
Log LUN size	136 GB
Total database size for performance testing	1080 GB
Percent storage capacity used by Exchange database**	99 percent

**Storage performance characteristics change based on the percentage utilization of the individual disks. Tests that use a small percentage of the storage (~25 percent) may exhibit reduced throughput if the storage capacity utilization is significantly increased beyond what is tested in this paper.

Primary Storage Hardware

Storage Connectivity (Fiber Channel, SAS,SATA, iSCSI)	SAS
Storage model and OS/firmware revision	StorageTek 2530 6.17.52.10 www.windowsservercatalog.com/item.aspx?itemId=53b769e5-acbd-4119-7d29-27a55a49306f
Storage cache	680 MB
Number of storage controllers	2
Number of storage ports	2
Maximum bandwidth of storage connectivity to host	18 Gb/s (up to six 3 Gb/s host connections per controller)
Switch type/model/firmware revision	N/A
HBA model and firmware	SG-XPCIE8SAS-E-Z01.18.00.00
Number HBAs/host	2
Host server type	Sun Microsystems, Inc. SunFire 4100 M2 four dual-core AMD Opteron 2216 processors 8129 MB RAM
Total number of disks tested in solution	24
Maximum number of spindles can be hosted in the storage	As configured: 24 Maximum configuration: 36

Primary Software Configuration

HBA driver	1.21.26.0
HBA QueueTarget Setting	N/A
HBA QueueDepth Setting	N/A
Multi-Pathing	SMIA-wsx64-01.01.32.04 MPIO DSM
Host OS	Microsoft Windows Server 2003,Enterprise Edition x64 SP2
ESE.dll file version	08.00.0685.024
Replication solution name/version	N/A

Primary Disk Database Configuration

Disk type, speed and firmware revision	146 GB 15,000 rpm 3 Gb/s-0791
Raw capacity per disk (GB)	136.7 GB
Number of physical disks in test	8
Total raw storage capacity (GB)	1094 GB
Disk slice size (GB)	N/A
Number of slices per LUN or number of disks per LUN	N/A
RAID level	RAID 10
Total formatted capacity	539 GB
Storage capacity utilization	49 percent
Database capacity utilization	33 percent

Primary Disk Log File Configuration

Disk type, speed and firmware revision	146 GB 15,000 rpm 3 Gb/s-079
Raw capacity per disk (GB)	136.7 GB
Number of Spindles in test	4
Total raw storage capacity (GB)	550.8 GB
Disk slice size (GB)	N/A
Number of slices per LUN or number of disks per LUN	N/A
RAID level	RAID 10
Total formatted capacity	135 GB

Best Practices

Exchange server is a disk-intensive application. Based on the testing run using the ESRP framework, we would recommend the following to improve the storage performance. For Exchange 2007 best practices on storage design, please visit <http://technet.microsoft.com/en-us/library/bb124518.aspx>

Core Storage

- Use the Microsoft Diskpart utility to align the sectors of all Exchange 2007 storage volumes before formatting. The Diskpart value should be set to 64.
- Do not share Exchange 2007 disks with any other applications that are I/O intensive. This will have a negative effect on your disk subsystem performance.
- When possible, separate the sequential writes of logs from the random access of database volumes.
- Primary focus should be on the speed of the disks and not size.

Contact for Additional Information

sun.com/service/serviceplans/index.jsp

Test Result Summary

This section provides a high level summary of the test data from ESRP and the link to the detailed html reports which are generated by ESRP testing framework.

Reliability

A number of tests in the framework are to check Reliability tests runs for 24 hours. The goal is to verify the storage can handle high IO load for a long period of time. Both log and database files will be analyzed for integrity after the stress test to ensure no database/log corruption. The following list provides an overview: (click on the underlined word will show the html report after the reliability tests run)

- No errors were reported in the saved eventlog file
- No errors were reported during the database and log checksum process

Primary Storage Performance Results

The Primary Storage performance testing is designed to exercise the storage with maximum sustainable Exchange type of IO for 2 hours. The test is to show how long it takes for the storage to respond to an IO under load. The data below is the sum of all of the logical disk I/Os and average of all the logical disks I/O latency in the two-hour test duration. Each server is listed separately and the aggregate numbers across all servers is listed as well.

Individual Server Metrics

The sum of I/Os across Storage Groups and the average latency across all Storage Groups on a per server basis.

Database I/O

Database I/O

Average database disk transfers/sec	2058.971
Average database disk reads/sec	1182.499
Average database disk writes/sec	876.471
Average database disk read latency (ms)	18
Average database disk write latency (ms)	14

Transaction Log I/O

Average log disk writes/sec	382.993
Average log disk write latency (ms)	1

Conclusion

This document is developed by storage solution providers, and reviewed by Microsoft Exchange Product team. The test results/data presented in this document is based on the tests introduced in the ESRP test framework. Customer should not quote the data directly for his/her pre-deployment verification. It is still necessary to go through the exercises to validate the storage design for a specific customer environment. ESRP program is not designed to be a benchmarking program; tests are not designed to getting the maximum throughput for a giving solution. Rather, it is focused on producing recommendations from vendors for Exchange application. So the data presented in this document should not be used for direct comparisons among the solutions.

