Intel RAID Sizing and Best Practices for Microsoft Exchange Server 2007

Intel® RAID Controller SRCSASJV
Intel® RAID Controller SRCSASRB
Intel® Server Platform S5000P

Contents

Reference Documentation 4
Hardware Components 4
Software Components 4
Introduction 5
I/O Types 6
RAID 7
RAID Definitions 7
RAID Controllers/Adapters 8
RAID Recommendations for Exchange Server 8
Disk Drive Types 10
Formatting and Alignment 11
Microsoft Exchange Server Jetstress 12
Exchange Server 2007 Configurations for Intel Server Solutions 14
Conclusion 19

www.intel.com/go/RAID

Reference Documentation

Body Text

Hardware Components

Quantity	Item	Manufacturer	Model
1	Intel® Server Board	Intel	S5000PSL. Additional information is available at: http://www.intel.com/products/server/motherboards/s5000psl/s5000psl-overview.htm.
1	Intel® Server Chassis	Intel	SC5400. Additional information is available at: http://www.intel.com/products/server/chassis/sc5400/sc5400-overview.htm.
1 (Select one)	Intel® RAID Controller	Intel	SRCSASJV. Additional information is available at: http://www.intel.com/products/server/raid-controllers/srcsasjv/srcsasjv-overview.htm.
	Intel® RAID Controller	Intel	SRCSASRB. Additional information is available at: http://www.intel.com/products/server/raid-controllers/srcsasrb/srcsasrb-overview.htm.
2	Intel® Xeon® Processors	Intel	Please refer to the Supported Processor List at http://www.intel.com/support/motherboards/server/sb/CS-022346.htm. Processors must support Intel EM64T.
4 GB minimum	Memory	Any supported	Please refer to the Tested Memory List at http://www.intel.com/support/motherboards/server/s5000psl/sb/CS-022924.htm.
1 (60 GB minimum)	SAS or SATA 3.5-inch hard drives	Any supported	Please refer to the Server Hard Drive Validation Test Report at http://www.intel.com/support/motherboards/server/sb/CS-025416.htm.

Table 1 - Intel® Server Board S5000PSL Hardware Configuration

Software Components

Item	Version	Manufacturer	Comment
1	Windows Server 2003 R2	Microsoft	Any 64-bit edition
1	Exchange Server 2007	Microsoft	Available in 64-bit only

Table 2 - Installation Software BOM

Introduction

This document provides basic storage sizing, configuration guidance and best practices for Exchange Server 2007 using an Intel® RAID Controller SRCSASJV or Intel® RAID Controller SRCSASRB on the Intel S5000 server motherboard family and the Intel SC5400 system chassis family.

Microsoft Exchange Server 2007 is a 64-bit application and must be installed on a 64-bit version of Windows Server. Transactional input/output (I/O) requirements for Exchange Server 2007 have been reduced from previous releases due to its 64-bit nature and increases in database cache and other improvements,.

Microsoft Exchange Server 2007 can be deployed using directly-attached disk storage using an Intel RAID controller card and supported disk drives installed into the disk drive cages of the Intel SC5400 chassis. The basic requirements for storage for Exchange Server 2007 are that there is enough capacity, acceptable disk latency and response time, and enough disk throughput to meet service level agreements. The disk storage must be formatted using NTFS.

This document is divided into several sections. These are:

- I/O Types
- RAID (including definitions and recommendations)
- Disk Drive Types
- Formatting and Alignment
- Microsoft Exchange Server Jetstress
- Exchange Server 2007 Configurations for Intel Server Solutions

The first few sections are background sections and are required reading in order to understand the reasons for the last section providing specific sizing and configuration guidance.

I/O Types

There are four basic types of I/O patterns:

- Random Read
- Random Write
- Sequential Read
- Sequential Write

Random I/O has no particular pattern or order to the I/O activity, such as with database transaction processing or email activity. In random I/O, data is accessed all over the storage system at what appears to be random locations.

Sequential I/O occurs in sequential order, such as with large batch processing, the writing of transaction logs and other similar activities. In sequential I/O, data is typically accessed in sequential order.

Microsoft Exchange Server performs both random and sequential I/O, but for two different reasons. General email activity such as reading and sending email generates random I/O, as different users read and send email of various sizes at different times, in essentially an unpredictable manner. Exchange generates sequential I/O when it writes to its transaction log, and if recovery is need, the transaction log is always read in a sequential manner.

Exchange Server 2007 reads and writes randomly to its database in 8K byte increments. The transaction logs are written sequentially in various sizes, depending on several factors including the relative load on the system.

The typical ratio of database reads to database writes in an Exchange Server 2007 database is approximately 1:1. This ratio was different in earlier versions of Exchange Server. For additional details on Exchange database read/write ratios, see: http://technet.microsoft.com/en-us/library/bb738147.aspx.

RAID

RAID Definitions

Redundant Array of Independent Disks (RAID) are designed to provide improved performance or reliability or both, when compared to single, individual disk drives. The major types of RAID are shown below, with their design characteristics.

RAID Type	Characteristics
RAID 0	Interleaving or "striping" data across two or more disks
RAID 1	Disk mirroring - identical data written on two different disks
RAID 5	Data striping across multiple disks with single parity
RAID 6	Data striping across multiple disks with double parity

Table 3 - RAID Types

RAID 0 provides improved performance only and does not provide any additional reliability features. The improved performance comes because the data is "striped" across all the disks in the RAID group, achieving parallel access across multiple disks simultaneously.

RAID 1 improves reliability and performance, and can be combined with the other types of RAID to make striped sets of mirrors. Read performance is improved because either disk in the mirror can be read and typically the controller will choose the least busy disk in the mirror for reads when possible. In the event of a disk drive failure, the other disk in the mirror is available for I/O operations. RAID 1 is more expensive than the other types of RAID because it requires more disk drives for an equivalent usable capacity.

RAID 5 and RAID 6 provide improved reliability and performance over individual disks by including rotating parity within the disk group, either single parity (RAID 5) or double parity (RAID 6). The performance gains are achieved similar to striping in RAID 0. In the event of a disk drive failure, the data can be reconstructed onto a hot spare disk. RAID 5 protects against the failure of one drive in the RAID group. RAID 6 protects against the failure of two drives in the RAID group. RAID 5 and RAID 6 perform well with read operations, but not as well with write operations because of the additional writes that are needed in the parity calculations.

The choice of RAID is determined by a combination of factors, including the desired performance, reliability and available budget. Different types of RAID provide different benefits, as shown below.

RAID Type	Transactional I/O Performance	Capacity Utilization	Disk Failure and Rebuild Performance
RAID 0	Good	Best	Not Applicable

RAID 1	Best	Least	Best
RAID 5	Good	Better	Good
RAID 6	Good	Good	Better
RAID 10	Best	Least	Best

Table 4 - RAID Benefits

In general, RAID 10 provides the best performance and availability. RAID 5 provides the best capacity utilization while achieving good performance and availability. Although RAID 5 and RAID 6 are viable choices for Exchange Server 2007 environments, compared to RAID 10, RAID 5 and RAID 6 are at a disadvantage due to the Exchange Server mailbox database random read/write ratio of approximately 1:1.

RAID Controllers/Adapters

RAID controllers provide the intelligence to perform the RAID functions described above in hardware and are the adapters that connect the storage devices to the host server. Many RAID controllers, including the Intel RAID Controllers described in this document, support both Serial Attached SCSI (SAS) and Serial ATA (SATA) interfaces, which are interfaces available for enterprise and desktop disk drives. These RAID controllers typically provide more than enough throughput to handle the I/O performance needs of an application, such as Exchange Server, when connected to multiple disk drives simultaneously. Many RAID controllers include battery-backed cache memory for even higher levels of performance. RAID controllers often perform automatic error correction and hot-spare support, and support a variety of operating systems.

Tests described in this document were run on RAID controllers with Writeback Cache and Read Ahead Cache enabled. This configuration is recommended and requires the use of a battery backup ;module for the RAID controller to protect data in flight through controller cache. Disk drive cache should be disabled as during a power outage it is not possible to protect data in flight through the drive cache.

RAID Recommendations for Exchange Server

In general, configure storage for applications servers such as Exchange Server with performance and availability as design criteria. Use more disks and faster disks for best performance.

The best practices for Exchange Server are:

- Consider performance before capacity and select the highest performing disk drives your budget allows.
- Mailbox databases and log files should be placed on different RAID sets, due to their completely different I/O patterns.

- Exchange transaction logs should be placed on the lowest latency RAID sets, which is typically RAID 10 with battery-backed write cache.
- If possible, place the mailbox databases on RAID 10 sets.
- Do not place data from other applications on the mailbox or log RAID groups.

Disk Drive Types

There are four general categories of spinning disk drives:

1. Enterprise
2. Desktop
3. Notebook
4. Consumer

Although these general categories of spinning disk drives can be identified, there is some overlap between the features of the drives in each of these categories as manufacturers respond to changing market needs.

Enterprise disk drives have higher rotation speeds such as 10K and 15K RPM and moderately large capacities. These drives are designed to be powered up and operating 24 hours per day, 7 days per week and typically have 5 - 7 year warranties. Enterprise drives are available with Fibre Channel, SAS and SCSI interfaces. Some newer enterprise-class drives have more recently appeared on the market with SATA interfaces. Among the four types of spinning disk drives, enterprise drives typically consume the most electric power.

Desktop disk drives have moderately high rotation speeds and very large capacities. Some desktop disk drives may be designed for something less than 24 hour per day operation. These drives typically have 3 - 5 year warranties. Desktop drives typically have SATA interfaces.

Notebook disk drives are designed for notebook and laptop computers and are designed for smaller size and lower power consumption.

Consumer disk drives are designed for consumer devices such as camcorders and MP3 players. These disk drives are designed for the smallest form factors and the lowest power consumption.

Of these four categories, only the enterprise class SAS and SATA disk drives are suitable for an Exchange Server environment.

Formatting and Alignment

Exchange Server databases and logs must be stored on NTFS partitions. The format allocation size, sometimes known as the cluster size, affects performance. Best practices are to use "Basic Disks" to create a single partition on a LUN for Exchange data.

Data type	Format Allocation Unit Size
Mailbox database	64K
Log	4K (or the default size)

Table 5 - Format Allocation Sizes

Partition misalignment is common. Aligning sectors to track boundaries can have performance benefits, depending on the storage. For Windows Server 2003 R2 and earlier, it is recommended to use the Diskpart.exe command-line utility to create properly aligned partitions. Always use the storage vendor's recommendation for the alignment setting. If this is unknown, use 64K as the setting. An example diskpart command for accomplishing this is:

- `create partition primary align=64`

Windows Server 2008 uses a default alignment setting of 1 MB, and does not need any special alignment to be performed by the user. Complete details on disk alignment for Exchange Server are available at: http://technet.microsoft.com/en-us/library/aa998219.aspx.

Additional details on formatting and alignment for Exchange Server are available at: http://technet.microsoft.com/en-us/library/bb738145.aspx.

Microsoft Exchange Server Jetstress

Jetstress is a disk-intensive tool that helps administrators verify the performance and stability of a disk subsystem that is intended for a production Microsoft Exchange Server environment. Jetstress works with the Exchange Server Database storage engine to simulate the Exchange Server I/O workloads for the Exchange database (mailboxes) and logs. Jetstress is designed to match the Exchange Server 2007 disk I/O ratios and patterns of burstiness.

In the mailbox test profile, the number of mailbox users are specified, along with the mailbox size and the Exchange I/O per second (IOPS) rate. These tests can run from 2 hours to 24 hours in specific intervals. Tests of longer than 6 hours are considered stress tests. The goal of Jetstress is to identify several performance metrics of the disk subsystem and determine if that particular subsystem, in its particular configuration, will provide the performance that Exchange Server needs.

Microsoft defines these basic classes of users of Exchange Server 2007 email systems:

User type (usage profile)	Send/receive per work day (approx. 50 KB size)	Database cache per user	Estimated IOPS per user
Light	5 sent / 20 received	2 MB	0.11
Medium	10 sent / 40 received	3.5 MB	0.18
Heavy	20 sent / 80 received	5 MB	0.32
Very heavy	30 sent / 120 received	5 MB	0.48
Extra heavy	40 sent / 160 received	5 MB	0.64

Table 6 - Exchange Server 2007 User Profiles

Note that the estimated IOPS per user are for Exchange Server 2007. Previous versions of Exchange Server required higher IOPS for the same workloads.

Jetstress produces some overall metrics and some metrics for each database and log volume. The overall metrics include:

- Target I/O per second (number of mailboxes multiplied by the IOPS setting)
- Achieved I/O per second (should exceed Target I/O by at least 5%)

The key metrics produced by Jetstress for each volume for the databases and logs are:

- Average disk seconds/read (database latencies should not exceed 0.020 secs.)
- Average disk seconds/write (log latencies should not exceed 0.010 secs.)
- Disk reads/second
- Disk writes/second
- Average disk bytes/write (logs only)

Memory in the mailbox server also affects I/O performance. In general, more memory in the mailbox server reduces the amount of disk I/O generated by Exchange Server 2007 due to caching, up to about 32 GB of RAM. Although Exchange Server can use more than 32 GB of RAM, amounts of RAM above 32 GB do not significantly improve disk I/O performance. The basic mailbox server memory recommendations are:

- Light users: 2 GB plus 2 MB per mailbox
- Medium users: 2 GB plus 3.5 MB per mailbox
- All Heavy users: 2 GB plus 5 MB per mailbox

Microsoft best practices for Exchange Server 2007 include keeping the individual databases at no more than 100 to 200 GB in size, depending on the replication method used. The individual databases are spread across Exchange Server "storage groups". As the mailbox size grows, Microsoft recommends to maintain the optimal size of each individual mailbox database file.

Exchange Server 2007 supports up to 50 storage groups. Adding a storage group increases the amount of database cache used for write activity, which has a positive impact on database write I/O. However, too many storage groups may reduce the amount of cache available for database reads, depending on the physical memory of the mailbox server. It is important to provide sufficient memory to include the mailbox load described above as well as the number of storage groups. For Exchange Server 2007 with SP1, the storage group memory requirements range from 2 GB for 1 - 4 storage groups, up to 15 GB for 49 - 50 storage groups.

The recommended configuration for the mailbox server role of a 1000-user mailbox of heavy users would require at least 7 GB of memory (2 GB plus 5 MB x 1000). This would allow for up to 20 storage groups. Additional memory is required if all the Exchange Server roles are to be run on the same server. The complete chart showing memory requirements for the Exchange 2007 server roles, mailboxes and storage groups is available in the "Exchange Server 2007 Planning Memory Configurations" page at http://technet.microsoft.com/en-us/library/bb738124.aspx.

Exchange Server 2007 Configurations for Intel Server Solutions

This section of the document provides background and guidance for sizing specific configurations of Exchange Server 2007 on the Intel platforms. This sizing includes different numbers of Exchange mailboxes using different storage configurations.

In order to verify the guidance provided in this document, a series of Exchange Server Jetstress tests were run with various 4, 6 and 10 disk drive configurations with the Intel SRCSASJV and SRCSASRB RAID controllers in the Intel SC5400 system chassis using an Intel S5000PSL server board. The server configuration was:

- Intel SC5400 system chassis with 4-drive and 6-drive cages, holding a maximum of 10 3.5-inch disk drives
- Intel S5000PSL server board
- Dual Intel Xeon E5320 processors, 1.86 GHz, 8 total cores
- 4 GB RAM
- 1 internal SATA 320 GB drive for the operating system and applications
- Windows Server 2003 R2 Enterprise x64 Edition

The two Intel RAID controllers, SRCSASJV and SRCSASRB, were tested by installing them, one at a time, into the PCI-Express 1.0 x8 slot of the SC5000PSL server board. These RAID controllers provide more than enough throughput for the configurations tested. The RAID controller should not be the bottleneck in terms of performance, and these controllers achieve that design goal.

Model	Devices supported	Drive type	Interface Speed	Ports	Cache Memory	PCI Express 1.0 Interface
SRCSASJV	240	SAS/SATA	3 Gb/sec	8 flexible	512 MB included Up to 1 GB	x8
SRCSASRB	32	SAS/SATA	3 Gb/sec	8 internal	256 MB	X4

Table 7 - Intel RAID Controller Specifications

Two sets of disk drives were tested in the 4, 6 and 10 drive configurations. These drives are shown in the table below.

Category	Model	Interface	Interface Speed	Cache Memory	Capacity	RPM
Enterprise SAS	Seagate Cheetah 15K.5	SAS	3 Gb/sec	16 MB	146 GB	15000
Enterprise SATA	Seagate Barracuda ES2	SATA	3 Gb/sec	32 MB	500 GB	7200

Table 8 - Tested Disk Drive Specifications

The following table provides the list of tested configurations that passed the Jetstress tests. All these tests were conducted using RAID 1 or RAID 10 in order to achieve the best performance. Some preliminary tests were run with RAID 5, but the results showed lower performance.

For these tests, all mailboxes were sized at 250 MB.

Configuration Name	Drive type	Number of drives	Database RAID configuration	Number of database drives	Log RAID configuration	Number of log drives
E-4	Enterprise	4	RAID 1	1 pair (2)	RAID 1	1 pair (2)
E-6	Enterprise	6	RAID 10	2 pairs (4)	RAID 1	1 pair (2)
E-10	Enterprise	10	RAID 10	4 pairs (8)	RAID 1	1 pair (2)
D-4	Desktop	4	RAID 1	1 pair (2)	RAID 1	1 pair (2)
D-6	Desktop	6	RAID 10	2 pairs (4)	RAID 1	1 pair (2)
D-10	Desktop	10	RAID 10	4 pairs (8)	RAID 1	1 pair (2)

Table 9 - Configuration Name Specifications

The performance data for the following tables refers to the respective configurations described above. Commentary on the performance data follows the tables below. The stress tests were run for a continuous 24-hour period, and were configured to meet the "very heavy" user workload level.

The test results below represent the high-end of the performance achievable with respect to the number of mailboxes of size 250 MB for each of the disk drive configurations and user usage profile. In many, but not all cases, doubling the size of the mailbox and reducing the number of mailboxes by half will achieve similar performance results.

Controller: SRCSASJV

Config. Name	Storage Groups	IOPS per user	Number of mailboxes	Available capacity	Database size	% capacity used	Target IOPS	Achieved IOPS	Thread Count
E-4	1	0.5	500	136 GB	122.1 GB	89%	250	268.313	4
E-6	2	0.5	900	272 GB	109.9 GB x 2	81%	450	541.255	8
E-10	4	0.5	1800	544 GB	109.9 GB x 4	81%	900	960.381	12
D-4	1	0.3	550	465 GB	134.3 GB	28%	165	168.423	3
D-6	3	0.3	800	930 GB	65.1 GB x 3	20%	240	304.335	3
D-10	3	0.3	1200	1860 GB	97.7 GB x 3	15%	360	392.385	4

Table 10 - SRCSASJV Jetstress Test Results

Controller: SRCSASRB

Config. Name	Storage Groups	IOPS per user	Number of mailboxes	Available capacity	Database size	% capacity used	Target IOPS	Achieved IOPS	Thread Count
E-4	1	0.5	500	136 GB	122.1 GB	89%	250	265.185	4
E-6	2	0.5	900	272 GB	109.9 GB x 2	81%	450	552.54	8
E-10	4	0.5	1800	544 GB	109.9 GB x 4	81%	900	977.816	12
D-4	1	0.3	550	465 GB	134.3 GB	28%	165	159.956	3
D-6	3	0.3	800	930 GB	65.1 GB x 3	20%	240	310.885	3
D-10	3	0.3	1200	1860 GB	97.7 GB x 3	15%	360	408.248	4

Table 11 - SRCSASRB Jetstress Test Results

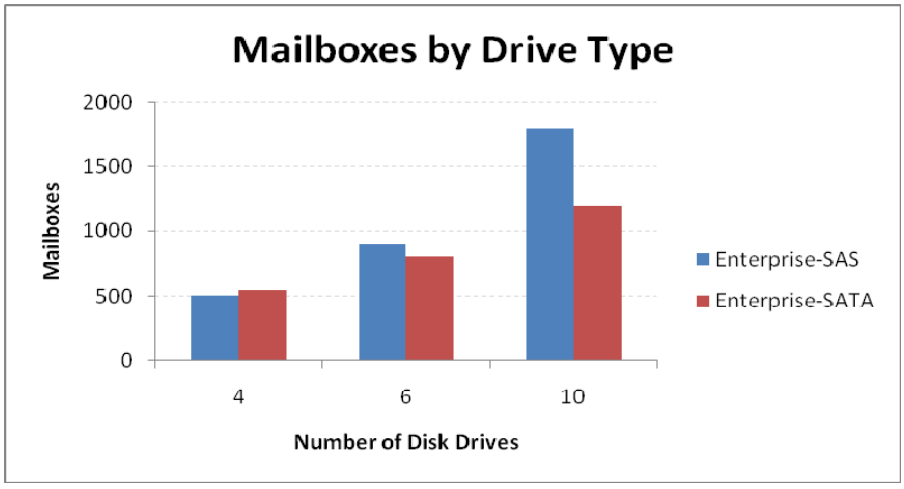

Table 12 - Mailboxes by Drive Type Chart

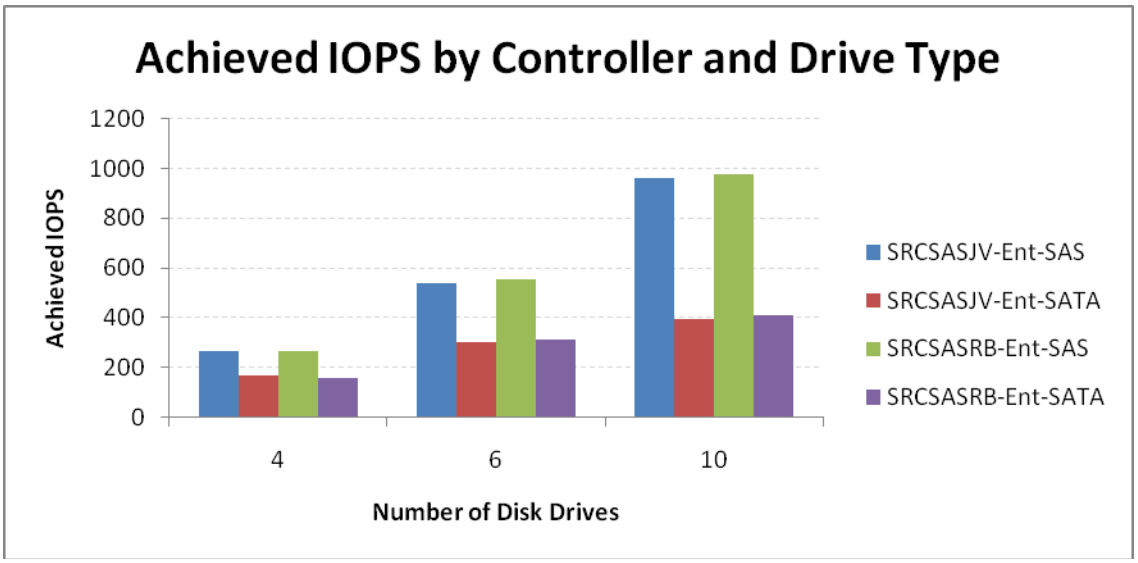

Table 13 - Achieved IOPS by Controller and Drive Type Chart

In the tests with these disk drive configurations, the two Intel RAID controllers, SRCSASJV and SRCSASRB, performed equally well. Both of these controllers are designed to handle larger numbers of disk drives than the Intel SC5400 server chassis will hold. It is important to insure that the RAID controller will handle more than the current workload, to allow room for growth. The SRCSASJV controller can also be connected to external SAS/SATA disk devices.

Although the enterprise and desktop disk drives are suitable for Exchange Server, as expected, the enterprise disk drives provided higher levels of performance than the desktop disk drives. The enterprise disk drives were able to satisfy the needs of the "very heavy" users for each of the mailbox configurations. The desktop drives were not able to sustain the necessary I/O rates for the heavy and very heavy users with similar numbers of mailboxes, but were able to satisfy the I/O rates of the light and medium users. The desktop drives were able to satisfy the I/O rates of very heavy users for a smaller number of mailboxes. If smaller numbers of mailboxes are needed, the desktop disk drives may provide adequate performance.

The enterprise disk drives were also able to handle higher thread counts. The thread count affects the number of disk transfers per second per storage group to the mailbox database and to the log file. As the thread count increases, the number of disk transfers per second for each storage group increases. Additional detail on the thread count is available at: http://technet.microsoft.com/en-us/library/bb643104.aspx.

Note that the desktop disk drives use relatively little of their total capacity for the Exchange Server databases, especially for the larger configurations. The desktop disk drives also provide significantly less performance than the enterprise disk drives. The unused space on the desktop disk drives is wasted.

Complete details of the results of these tests are available on the Microsoft Exchange Solution Reviewed Program (ESRP) web page.

Based on these Exchange Server workload tests, the following general statements can be made:

- Enterprise SAS disk drives will run out of capacity before they run out of performance.
- Enterprise SATA disk drives will run out of performance before they run out of capacity.

Conclusion

The Intel S5000PSL server, SC5400 chassis and RAID controllers described in this document provide excellent Exchange Server 2007 performance for organizations with up to as many as 1800 email users, which is more than enough for most small and medium-sized businesses. Both the SRCSASJV and SRCSASRB RAID controllers provide excellent performance and are good choices to control storage used for Exchange Server 2007.

Storage recommendations:

- Both SRCASJV and SRCSASRB provide sufficient performance for the Exchange Server 2007 targets defined herein. The SRCSASJV controller is recommended if additional scalability such as external storage is needed.
- Choose enterprise SAS disk drives to handle the heaviest of email workloads.
- Choose enterprise SATA drives to handle lighter email workloads or for smaller numbers of email users.

Memory recommendations:

- Memory should be configured according to the guidance provided for the number of email users and the Exchange server roles. A system with 8 - 16 GB RAM should be adequate for many small and medium-sized businesses who want to run all the Exchange Server 2007 roles on a single server.