



Dell EqualLogic Best Practices Series

# Sizing Microsoft Exchange 2010 on Dell EqualLogic PS6100 and PS4100 Series Arrays on VMware vSphere 5

A Dell Technical Whitepaper

This document has been archived and will no longer be maintained or updated. For more information go to the [Storage Solutions Technical Documents page on Dell TechCenter](#) or contact support.

**Storage Infrastructure and Solutions Engineering**

Dell Product Group

May 2012

THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.

© 2012 Dell Inc. All rights reserved. Reproduction of this material in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell.

Dell, the DELL logo, and the DELL badge, PowerConnect™, EqualLogic™, PowerEdge™ and PowerVault™ are trademarks of Dell Inc. Broadcom® is a registered trademark of Broadcom Corporation. Intel® is a registered trademark of Intel Corporation in the U.S. and other countries. Microsoft®, Windows®, Windows Server®, and Active Directory® are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries.

# Table of Contents

|            |  |    |
|------------|--|----|
| 1          | Introduction .....   | 2  |
| 1.1        | Purpose and Scope .....                                    | 2  |
| 1.2        | Target audience .....                                      | 2  |
| 1.3        | Terminology .....  | 2  |
| 2          | EqualLogic 6100/4100 product overview .....                | 3  |
| 3          | Microsoft Exchange and storage subsystem .....             | 4  |
| 3.1        | Configuration factors to consider .....                    | 5  |
| 3.2        | Considerations around Exchange DAG differentials.....      | 6  |
| 4          | Test topology and architecture overview .....              | 7  |
| 4.1        | Functional system design .....                             | 7  |
| 4.2        | Physical system configuration .....                        | 8  |
| 4.3        | Storage layout .....                                       | 9  |
| 5          | Validate the trend of each factor.....                     | 10 |
| 5.1        | Characterize the impact of user profile workload .....     | 11 |
| 5.2        | Characterize the mailbox size.....                         | 15 |
| 5.3        | Study the database volumes layout .....                    | 17 |
| 5.4        | Study the DAG number of database copies differential ..... | 20 |
| 5.5        | Characterize the RAID policy .....                         | 22 |
| 5.6        | Scaling up the number of users.....                        | 25 |
| 5.7        | Scaling out the SAN .....                                  | 30 |
| 5.8        | Assess the 6100 and 4100 family models.....                | 34 |
| 6          | Capacity planning and sizing .....                         | 38 |
| 6.1        | Capacity considerations.....                               | 38 |
| 6.2        | Performance considerations .....                           | 40 |
| 6.3        | Sizing example .....                                       | 41 |
| 7          | Best practices recommendations .....                       | 44 |
| Appendix A | Configuration details .....                                | 47 |
| A.1        | Hardware components .....                                  | 47 |
| A.2        | Software components.....                                   | 48 |
| A.3        | Network configuration details .....                        | 49 |
| A.4        | Host hypervisor and virtual machines configuration .....   | 51 |
| A.4.1      | Virtual network configuration .....                        | 51 |
| A.4.2      | Virtual Machines network configuration.....                | 54 |
| Appendix B | Microsoft Jetstress considerations.....                    | 55 |

# Acknowledgements

This whitepaper was produced by the PG Storage Infrastructure and Solutions of Dell Inc.

The team that created this whitepaper:

**Danilo Feroce, Puneet Dhawan, Suresh Jasrasaria, and Camille Daily**

We would like to thank the following Dell team members for providing significant support during development and review:

**Mark Welker**

## Feedback

We encourage readers of this publication to provide feedback on the quality and usefulness of this information by sending an email to [SISfeedback@Dell.com](mailto:SISfeedback@Dell.com).



[SISfeedback@Dell.com](mailto:SISfeedback@Dell.com)

# Executive Summary

The endless growth of storage capacity demand is showing no signs of waning. Disk drives with increasingly larger capacity are becoming available every year and the cost per gigabyte is routinely declining, and has progressed to such proportions that the terabyte is the new unit of storage capacity. A solution that provisions such large storage capacity, however, is not as simple as it appears. Applications and storage infrastructures must support such growth with the same or even improved performance to user data access. End users would not accept any degradation in the usability of their business tools solely because hundreds to thousands times more data is being accessed than in the past.

Microsoft® Exchange Server 2010 is able to address this storage growth challenge through its ability to implement and manage very large user mailboxes while reducing the storage access footprint due to significant redesigns in its database storage architecture. The recent introduction of the updated Dell™ EqualLogic™ PS Series family arrays (6100 and 4100) substantially raises the bar of both performance and capacity kindling many organizations to investigate the benefits of using them to power their corporate applications and storage-based services such as Microsoft Exchange Server.

Combining the messaging services and SAN platforms is a winning scenario for the end user expectations, but it necessitates the careful examination of IT professionals when it comes to planning and deploying a production environment. Understanding the current access footprint of the end users and services is a key element to properly begin the journey of sizing the right solution. Many additional variables, originating from either software or hardware, might influence the outcome in a powerful and surprising way if not weighted in advance.

We evaluated most of the variables surrounding the Exchange 2010 deployment decisions against the storage and identified some key findings summarized below.

- Exchange database maintenance and service tasks could have a considerable impact on the overall footprint of storage IOPS. The absolute value of their load is approximately steady, therefore when examined as a percentage, the value can profoundly influence environments with light load while it appears mitigated when compared instead with heavy workloads.
- Distributing a given set of user mailboxes with a known workload in fewer databases would greatly benefit the overall performance of the storage subsystem. Reproducible tests showed an application latency improvement between 55% and 70% when shifting from 20 to 2 deployed databases to support the same number of users. Nevertheless the flexibility of administration and protection of a large database must be weighed with caution.
- Database availability group (DAG) technology manifests a light, but consistent increment in generated IOPS due to the supplementary read activity against database log files. The number of replica copies is a factor of the increment moving from 2% overhead of the first copy, to 1% for each subsequent duplication of the number of copies.
- Traditional RAID policies comparison demonstrates how RAID 10 still performs ahead of other implementations. Around 5% from the closest competitor RAID 5, followed by RAID 50, and then RAID 6. Due to the EqualLogic design, RAID 5 and RAID 6 benefit from a greater number of active spindles. Performance is nevertheless close enough to persuade a decision maker to base his disk protection choice on capacity or on higher tolerance to failure characteristics, instead of pure performance.

# 1 Introduction

Microsoft Exchange Server is a versatile messaging solution heavily relying on the storage subsystem. The 2010 version provides access to more data per mailbox with same or improved efficiency.

The EqualLogic updated PS Series families (6100 and 4100), while maintaining the same building block approach and easy management, accommodate wider capacity needs and dispense increased performance.

Virtualization technologies and private clouds allow a more beneficial exploitation of hardware investments, new or preferable ways to provide high availability and reliability to IT solutions, and basically a better control of the expenditures.

Virtualization is also often associated with an enhanced chance of curtailment of the datacenter physical footprint. When combined with server blades architecture, it exhibits an unparalleled processing density to the corporations, eager to gauge and control their energy consumption and ultimately their total cost of ownership (TCO).

A proper amalgamation of these four main pillars (storage, blade servers, virtualization, and Microsoft Exchange application) requires an expanded canvass to address the additional variables that might influence the outcome of a solution.

## 1.1 Purpose and Scope

This white paper presents the results of a study centered on the characterization of I/O patterns of Microsoft Exchange Server when deployed on a storage subsystem based on EqualLogic SAN and in a dense server environment built on Dell™ PowerEdge™ blade servers. It will guide Exchange and SAN administrators to understand their messaging workload, and predict their SAN size requirements. The scope of this paper is a virtual infrastructure built on VMware® vSphere® connecting to the SAN directly from the virtual machines and storage sizing, leaving aside server sizing activity.

## 1.2 Target audience

This paper is primarily intended for IT professionals (IT managers, Solution Architects, Exchange and Storage Administrators, and System and VMware Engineers) who are involved in defining, planning and/or implementing Microsoft Exchange Server 2010 infrastructures and would like to investigate the benefits of using EqualLogic storage. This document assumes the reader is familiar with Microsoft Exchange functionalities, EqualLogic SAN operation, and VMware system administration.

## 1.3 Terminology

The following terms will be used throughout this document.

**Group:** Consists of one or more EqualLogic PS Series arrays connected to an IP network that work together to provide SAN resources to host servers.

**Member:** Identifies a single physical EqualLogic array.

**Pool:** A logical collection that each member (array) is assigned to after being added to a group and contributes its storage space to the entire pool.

**Hypervisor:** Denotes the software layer in charge of managing the access to the hardware resources, sitting above the hardware, and in between the operating systems running as 'Guests'.

**Virtual Machine:** Denotes an operating system implemented on a software representation of hardware resources (processor, memory, storage, network, etc.). Virtual machines are usually identified as 'guests' in relation with the 'host' operating system that execute the processes to allow them to run directly on the hardware.

**Exchange DAG:** Database Availability Group is a pool of networked Exchange mailbox servers that hosts multiple copies of the same Exchange databases.

**B-Tree (balanced tree):** a tree data structure where a node can have a variable number of child nodes, commonly used in databases to maintain data sorted in a hierarchical arrangement. It allows efficient data access to the pages for insert, delete, and searches.

## 2 EqualLogic 6100/4100 product overview

Dell™ EqualLogic™ PS6100 Series arrays are designed to meet the performance and availability needs of application and virtualization environments in medium to large enterprises. These virtualized iSCSI storage area networks (SANs) combine intelligence and automation with fault tolerance to provide simplified administration, rapid deployment, enterprise performance and reliability, and seamless scalability using innovative Fluid Data™ technology.

The Dell EqualLogic PS4100 series is designed to address the needs of remote office/branch office (RO/BO) and small-to-medium business storage deployments. Remote offices receive the benefits of data center class IT solutions, such as storage consolidation and virtualization, at an affordable entry point.

PS6100 arrays fit your needs with 2.5" or 3.5" drives in 2U or 4U form factors—all while delivering an increase in drives per array of up to 50% over the previous generation. With an option of all SSD drives, a mix of SSD and 10K SAS drives in a single array, all 10K SAS or all 15K SAS drives in a single array, EqualLogic PS6100/4100 series array models offer the flexibility in capacity and performance to best match application needs.

### 3 Microsoft Exchange and storage subsystem

Microsoft Exchange Server product has a diversified set of components and services working together to accomplish the goal of supporting the most dissimilar requirements to design and deploy a messaging infrastructure within an organization. From a storage perspective, the most relevant role performed in Exchange infrastructure is the mailbox server role, since it ultimately governs the retrieval, storage, and availability of users data for the rest of the infrastructure in order to have it routed, presented, etc.

Appropriate sizing of the mailbox role servers in an organization is the primary best practice to avoid poor performance issues, or administrative overhead to redesign the deployment layout to adapt to new or changed users requirements.

We should follow the hierarchy of Microsoft Exchange Server 2010 underlying logical components to understand the interaction between the Exchange mailbox server role and the storage subsystem. The access to mailbox databases (or public folder database, when implemented) is the key element that primarily generates I/O to the storage subsystem. However, while a database is a logical representation of a collection of user mailboxes, it is at the same time an aggregation of files on the disk which are accessed and manipulated by a set of Exchange services (i.e. Exchange Information Store, Exchange Search Indexer, Exchange Replication Service, Microsoft Exchange server) following a different set of rules.

**Database file** (\*.edb) is the container of users mailbox data. Its content, broken into database pages (32KB in size), is primarily read and written in a random fashion as required by the Exchange services running on the mailbox server role. A database has a 1:1 ratio with its own .edb database file. The maximum supported database size in Exchange Server 2010 is 16 TB, where the Microsoft guidelines recommend a maximum 200 GB database file in a standalone configuration and 2 TB if the database participates in a replicated DAG environment.

**Transaction Logs** (\*.log) are the container where all the transactions that occur on the database (create, modify, delete messages, etc.) are recorded. Each database owns a set of logs, and keeps a 1 to many ratio with them. The logs are written to the disk sequentially, appending the content to the file. The logs are read only when in a replicated database configuration (DAG) or in the event of a recovery.

**Checkpoint file** (\*.chk) is a container for metadata indicating when the last flush of data from the memory cache to the database occurred. It has a very limited size (8KB), and although repeatedly accessed its overall amount of I/O is limited. The database keeps a 1:1 ratio with its own checkpoint file and positions it in the same folder location as the log files.

**Content Index files** (catalog folder, multiple file extensions) are flat files, representing the Search Catalog, built by the Microsoft Search Service. The client applications connected to Exchange Server benefit from this catalog by performing faster searches based on indexes instead of full scans.

Microsoft Exchange Server uses a proprietary format named Extensible Storage Engine (ESE) to access, manipulate, and save data to its own mailbox databases. The same format is now employed on the Exchange HUB server role for queue databases. The ESE technology, previously known as Jet Database Engine, has evolved through several versions of Exchange Server releases and created different ramifications within the Microsoft products since its inception (Microsoft Access, Active Directory, File Replication Service, WINS server, Certificate Services). The ESE is an Indexed Sequential



Access Method (ISAM) technology, which organizes the database data in B-Tree structures, and as such the databases are populated with the effort of keeping the data together or adjacent. Considering this event does not always occur. Such structured databases benefit from external tasks directed towards the reorganization or defragmentation of the database data itself to restore the optimal data contiguity.

**Note:** For additional information about the Exchange 2010 Store refer to Microsoft documentation *Understanding the Exchange 2010 Store*, available at: <http://technet.microsoft.com/en-us/library/bb331958.aspx>

To summarize, an Exchange 2010 database itself is subject to a subset of tasks producing storage access:

- the regular read and write access required to retrieve user mailboxes data and store it (according to the Exchange cache policy);
- the online defragmentation and compacting activities due to the B-Tree optimization
- the database maintenance, which includes dumpster cleanup, deleted mailboxes purge, and other activities addressing logical objects support; and
- the database scanning checksum to verify data blocks integrity (sequential read activity), which can be set as a background 24x7 activity or in a scheduled time window.

Additionally, Exchange Server offers a specialized offline defragmentation task that can only be performed manually, taking advantage of the ESEUTIL.EXE command line tool, while the database is dismounted. The principal goal of this task is to reclaim the empty space left in a database by the online defragmentation, shrinking the size of the .edb file itself, and thus returning the free space back to the operating system volume.

Any further discussion around offline defragmentation has been purposely omitted from this paper as this task does not contribute toward the scope of our study. Furthermore, it is not recommended to include offline defragmentation in a regular maintenance plan due to the disruption in the availability of the database, the rupture of the logs chain, and the obligation of database re-seeding in case of Database Availability Group configuration.

### 3.1 Configuration factors to consider

After determining the different I/O activities generated against the storage subsystem for a single database, the factors that influence the overall I/O footprint of an entire Exchange mailbox server role and the several variables of a deployment were taken into consideration.

**Mailbox usage profile** denotes the usage characteristics of a mailbox (i.e. send, receive, open or delete items). It is commonly defined by the amount of messages sent and received per day and the average message size. It is also translated into transactional IOPS per mailbox through considerations made around the size of database cache allocated per each mailbox.

**Mailbox size** quantifies the maximum space a mailbox will be allowed to grow, enforced by a quota policy, or more generally the average mailbox size in a corporate messaging environment. It primarily affects the capacity requirement of a mailbox role server and the planned size of the databases hosted by the server. Moreover it influences the IOPS response due to the wider physical disk surface that must be accessed to retrieve or store the data.

**Number of Databases** designates the database layout and mailbox distribution across the databases. Each Exchange database is managed as single administrative unit and is served by a set of services with a 1:1 ratio (defragmentation, maintenance, logs generation).

**Database and Log placement** indicates whether the .edb and log files reside on the same volume or are deployed on isolated volumes. The historical requirement to split database and logs into different volumes, or rather physical drives, arose from two different reasons:

- Performance: the different I/O pattern of these two streams of data (random reads/writes versus sequential writes) and the aim to associate them with the most fitting storage device (i.e. rotational speed, RAID level).
- Reliability: the contemporaneous loss of both the database and logs could jeopardize the recoverability of users data depending on the backup methodology used.

The Exchange Server 2010 I/O footprint reduction, when compared with former versions, nullifies the perceived performance issue reported above. In addition, the use of a SAN as a storage subsystem provides different approaches to the data protection of Exchange mailbox server role that does not require the volume split mentioned above (i.e. EqualLogic Smart Copies). For additional references refer to the white paper mentioned in the notes box at the end of this section.

**Mailbox count** represents the number of user mailboxes hosted by the mailbox role server. With the increment of this value the amount of IOPS, capacity allocated and logs generation increase. Furthermore an elevated number of mailboxes hosted by a single server intensify the demand for a highly available solution

**High availability footprint** refers to the overhead of having a highly available solution in place with data replication involved. See the section titled, "Considerations around Exchange DAG differentials".

**Data protection footprint** identifies the planned amount of IOPS spent by the solution established to protect the mailbox database data. Since a customized solution is usually tailored to each distinct production environment (for example 24x7, or 9 to 5) the impact is greatly variable and depends upon the degree of parallelism between the regular data access and the data protection access. Where examining the aspects of protecting Exchange mailbox server role data requires a more comprehensive explanation, refer to the white paper mentioned in the notes box below.

For additional information about Exchange data protection options with Dell EqualLogic SAN refer to *Best Practices for Enhancing Microsoft Exchange Server 2010 Data Protection and Availability using Dell EqualLogic Snapshots*, available at:  
<http://www.delltechcenter.com/page/Enhancing+Microsoft+Exchange+Server+2010+Data+Protection+and+Availability+with+EqualLogic+Snapshots>

### 3.2 Considerations around Exchange DAG differentials

A Database Availability Group (DAG) is a pool of up to 16 networked servers that hosts multiple copies of the same Exchange database or databases where only one of the copies is active at a specific point in time within the group; the other copies are passive and contain data sourced from replicated and replayed transaction logs. While implemented, it directs some deviations in the storage access patterns and in the Exchange memory cache behavior.

**Transaction Logs access** is now differentiated between the conventional sequential writes pattern, and by an additional sequential read access, required to perform the replication activities.

**Log Checkpoint depth** refers to the amount of logs written to the disk and containing transactions not yet flushed to the database file. It is usually set at 20 in a standalone server install, and increased to 100 in a DAG configuration. The outcome of this change is a contraction in the write I/O for the given database, since the opportunity to combine user activity changes in memory (coalescing), and thus to reduce I/O increases.

## 4 Test topology and architecture overview

The findings presented in this paper are a result of testing conducted on a Microsoft Windows infrastructure, built on VMware vSphere 5.0 hypervisor, and accessing storage on EqualLogic SAN. We took advantage of the Microsoft Exchange Jetstress simulation tool in order to simulate the different workloads specified in section 5.

Our test architecture consisted of a core building block of one Exchange mailbox role server paired with one storage array. This block was then scaled horizontally to achieve the targeted workload level by multiplying the number of servers and arrays while firmly retaining the 1:1 ratio between them.

### 4.1 Functional system design

The functional elements of the test infrastructure are shown in Figure 1. Some key elements of the design were:

- Single Active Directory forest, single domain, single site (not strictly required for the tests),
- Centralized management and monitoring with dedicated resources, and
- Building block design approach for mailbox role servers.

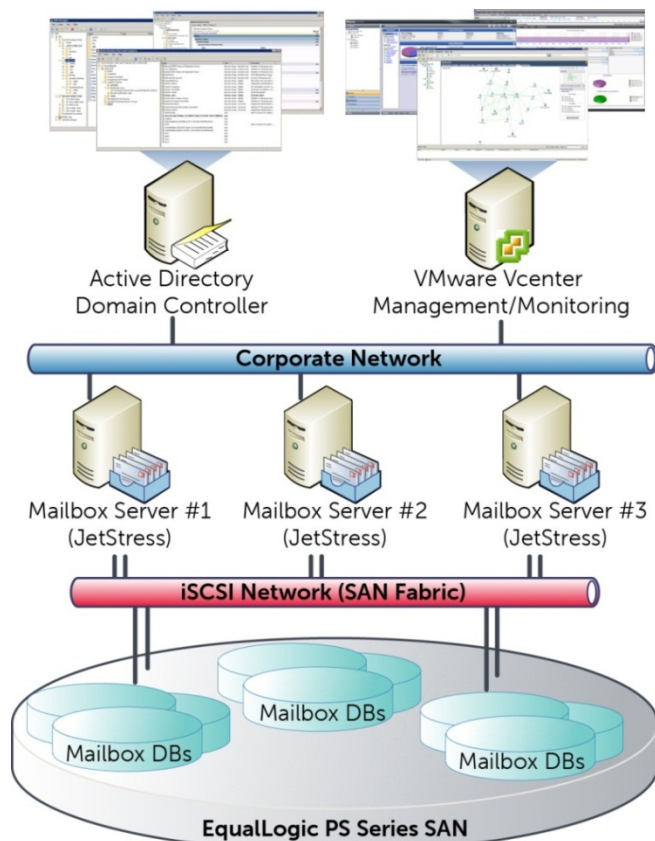


Figure 1 Functional system design diagram

## 4.2 Physical system configuration

The physical components of the test infrastructure were laid out as show in Figure 2.

We deployed this solution on Dell Blade servers planning for a greater datacenter density and flexibility of the solution. Some key aspects of the physical deployment were:

- Single M100E Blade enclosure with redundant management controllers (CMC) and fully populated power supplies configured for redundancy,
- Single EqualLogic iSCSI SAN provisioned alternatively by a single unit of one of the following array models: PS6100 3.5", PS6100X, PS6100E, PS4100XV 3.5", or PS4100E, and
- Single EqualLogic iSCSI SAN provided by up to 3 PS6100XV 3.5" arrays for the exercise to scale the SAN to multiple arrays.

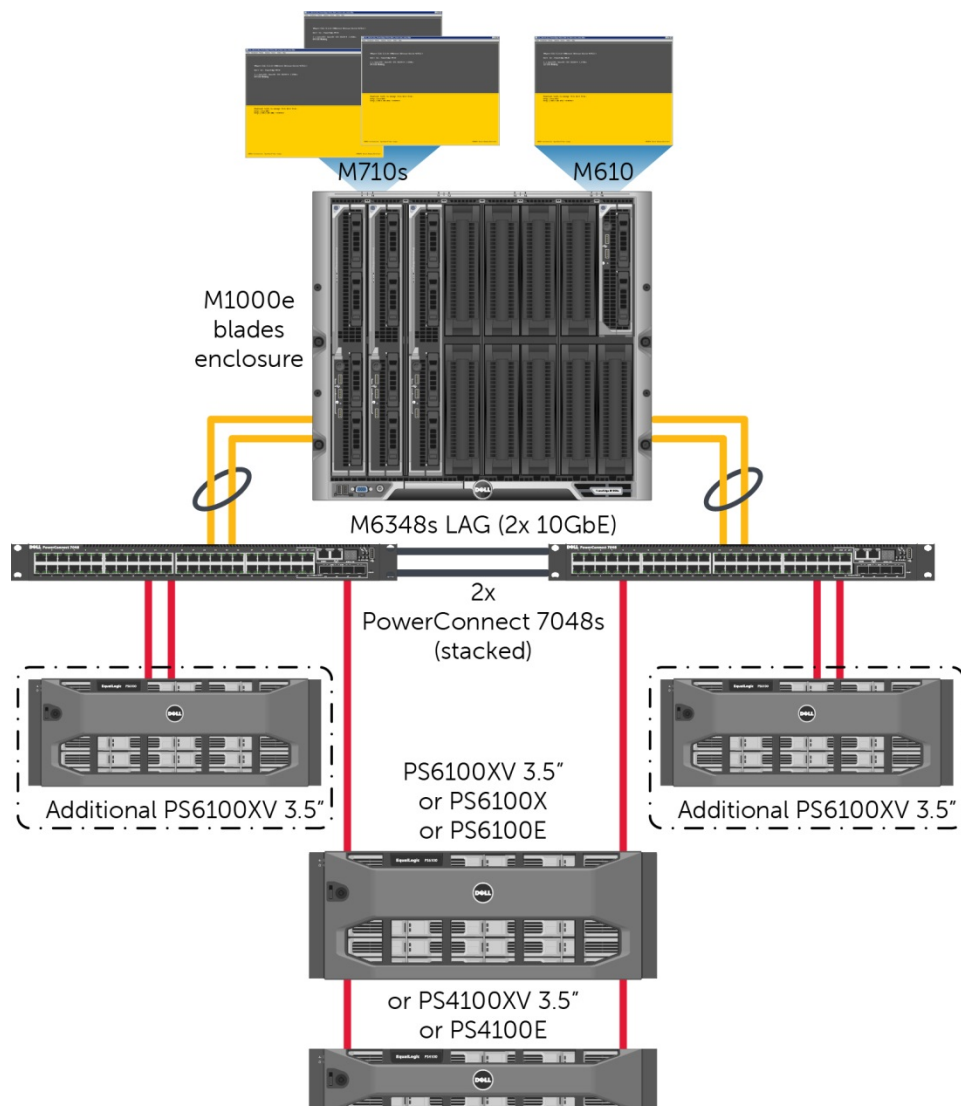


Figure 2 Physical system design diagram

- Dual PowerConnect M6220 Ethernet switches (stacked) to support LAN IP traffic (Fabric A)
- Dual PowerConnect M6348 Ethernet switches (stacked) to support the iSCSI data storage traffic on the server side (Fabric B),

- Dual PowerConnect 7048R Ethernet switches (stacked) to support the iSCSI data storage traffic on the SAN side, and
- Link Aggregation Group (LAG) consisting of 4 fiber connections (2 from each switch) between the Fabric B M6348s and the top of the rack 7048s.

More details of the test configuration setup, including a hardware and software list, SAN array characteristics, hypervisor and virtual machines relationship, network connections, and blade switch fabric paths are provided in Appendix A.

### 4.3 Storage layout

The EqualLogic SAN arrays and the volume underlying the Exchange databases were setup as:

- One EqualLogic group configured alternatively with one unit of each array model previously listed in the section titled, "Physical system configuration".
- One EqualLogic group configured with one to three array members for the simulation to scale the SAN to multiple arrays.
- One storage pool defined within the group and including the single member or all the members of the group for the test case of multiple arrays.
- RAID policy dependent and specified as part of each test case.
- Five volumes created in the pool, unless specified in the test case. One volume dedicated to each Exchange mailbox database and a set of file logs with a 1:1 ratio.

Figure 3 reflects the common volume layout implemented on the EqualLogic SAN.

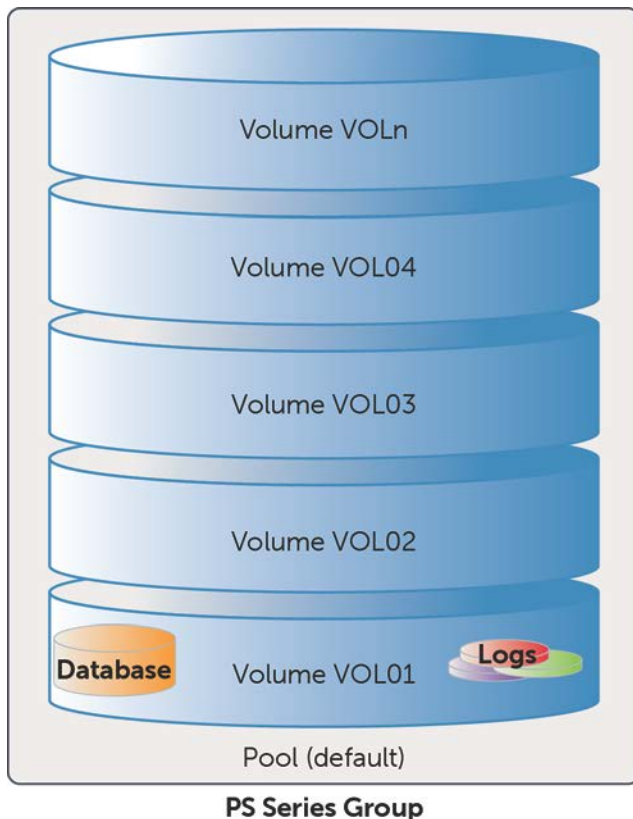


Figure 3 Volume and database/logs layout

## 5 Validate the trend of each factor

This set of tests was run to evaluate the storage performance trends of an EqualLogic SAN under the load of a reference Exchange mailbox role server configuration while the deployment factors were altered.

In order to predict storage needs for current or future implementations, an Exchange administrator and storage counterpart should be able to address the following quandaries:

- What if the usage profile of an average mailbox user in our organization changes?
- How much should an average mailbox be allowed to grow?
- Across how many mailbox databases should user mailboxes be distributed?
- What storage impact should be expected when DAG technology is adopted to increase mailbox availability?
- Does the RAID policy selected become a heavy constraint against Exchange performance?
- Which EqualLogic array model will be a satisfactory fit for the overall footprint of our mailboxes?
- How many users should be provisioned per each building block unit of deployment?
- Does the SAN scale horizontally following the same pace of growth of the workload?

The reference Exchange mailbox server role configuration used is reported in Table 1. Each test described in the following paragraphs explicitly reports variations made to this reference baseline.

For details about the simulation tool, Microsoft Jetstress 2010, refer to Appendix B.

**Table 1 Reference configuration**

| <b>Reference configuration: factors under study</b>                    |  |
|--|--|
| Number of simulated mailboxes/users                                    | 5,000 concurrent users   |
| Mailbox size   | 512MB  |
| Number of databases  | 5 databases (active)   |
| Databases size   | 500GB each   |
| Mailbox allocation   | 1,000 mailboxes per each mailbox database  |
| Number of database replica copies                                      | 1 (standalone)   |
| IOPS per mailbox / messages per day per mailbox                        | 0.18 IOPS / 150 messages   |
| RAID policy  | RAID 50  |
| Array model, amount of units, SAN configuration                        | 1x PS6100XV 3.5", one single pool (default)  |
| <b>Reference configuration: unaltered factors across the test sets</b> |  |
| Volumes  | One volume for each DB and associated Logs<br>Volume size equal 120% database size |
| Windows Disk/Partition, File System                                    | Basic disk, GPT partition, default alignment<br>NTFS, 64KB allocation unit size    |
| Background database maintenance  | Enabled  |
| Test duration  | 2 hours + time required to end DB checksum   |

Below is a list of metrics and pass/fail criteria recorded during testing. Most of this information is outlined by the Jetstress tool and the remainder is verified through Dell EqualLogic SAN Headquarters. Microsoft indications around thresholds for storage validation are reported as well.

**Database Reads Average Latency (msec)** is the average length in time to wait for a database read operation. It should be less than 20milliseconds (random reads according to Microsoft threshold criteria).

**Database Writes Average Latency (msec)** is the average length in time to wait for a database write operation. It should be less than 20milliseconds (random writes according to Microsoft threshold criteria).

**Logs Writes Average Latency (msec)** is the average length in time to wait for a log file write operation. It should be less than 10milliseconds (sequential writes according to Microsoft threshold criteria).

**Planned Transactional IOPS** are the target amount of IOPS for the test (calculated by multiplying the number of users by the IOPS per mailbox).

**Achieved Transactional IOPS** are the amount of IOPS really performed by the storage subsystem to address the transactional requests. The result should be not less than 5% of the planned IOPS to be considered a successful test iteration according to Microsoft Jetstress.

**LOGS IOPS** are the IOPS performed against the log files during the transactional test. They are not directly taken into account as part of the transactional IOPS, but tracked separately instead.

**Differential IOPS** are the IOPS generated for the DB maintenance and all the remaining activities on the storage subsystem, calculated as difference between the IOPS provisioned by the EqualLogic SAN and the previously reported transactional and logs IOPS.

## 5.1 Characterize the impact of user profile workload

The goal of the baseline workload analysis was to establish the storage trend, IOPS ratios, and relationship when applying the reference user workload and then varying it, while maintaining the remaining factors. The configuration parameters for the test are shown in Table 2.

**Table 2 Test parameters: variable workload**

| Reference configuration: variable factor        |   |
|---|---|
| IOPS per mailbox / messages per day per mailbox | 0.06 IOPS / 50 messages<br>0.18 IOPS / 150 messages<br>0.36 IOPS / 300 messages |
| Reference configuration: unchanged factors      |   |
| Number of simulated mailboxes/users             | 5,000 concurrent users  |
| Mailbox size                                    | 512MB   |
| Number of databases                             | 5 databases (active)  |
| Databases size                                  | 500GB each  |
| Mailbox allocation                              | 1,000 mailboxes per each mailbox database                                       |
| Number of database replica copies               | 1 (standalone)  |
| RAID policy                                     | RAID 50   |
| Array model, amount of units, SAN configuration | 1x PS6100XV 3.5", one single pool (default)                                     |



Exchange messaging service end users retrieve and store their data on the server databases traversing the Exchange Server services. In turn, the Exchange Information Store service running on the mailbox server benefits from the database cache to perform its own storage access instructions via the ESE interface. The database cache is retained in memory, thus the access to and from it is considerably faster than if performed with the storage subsystem directly. Furthermore, Exchange utilizes internal algorithms to combine changes to the same blocks of data before flushing them to the disks, achieving an even lower amount of IO access.

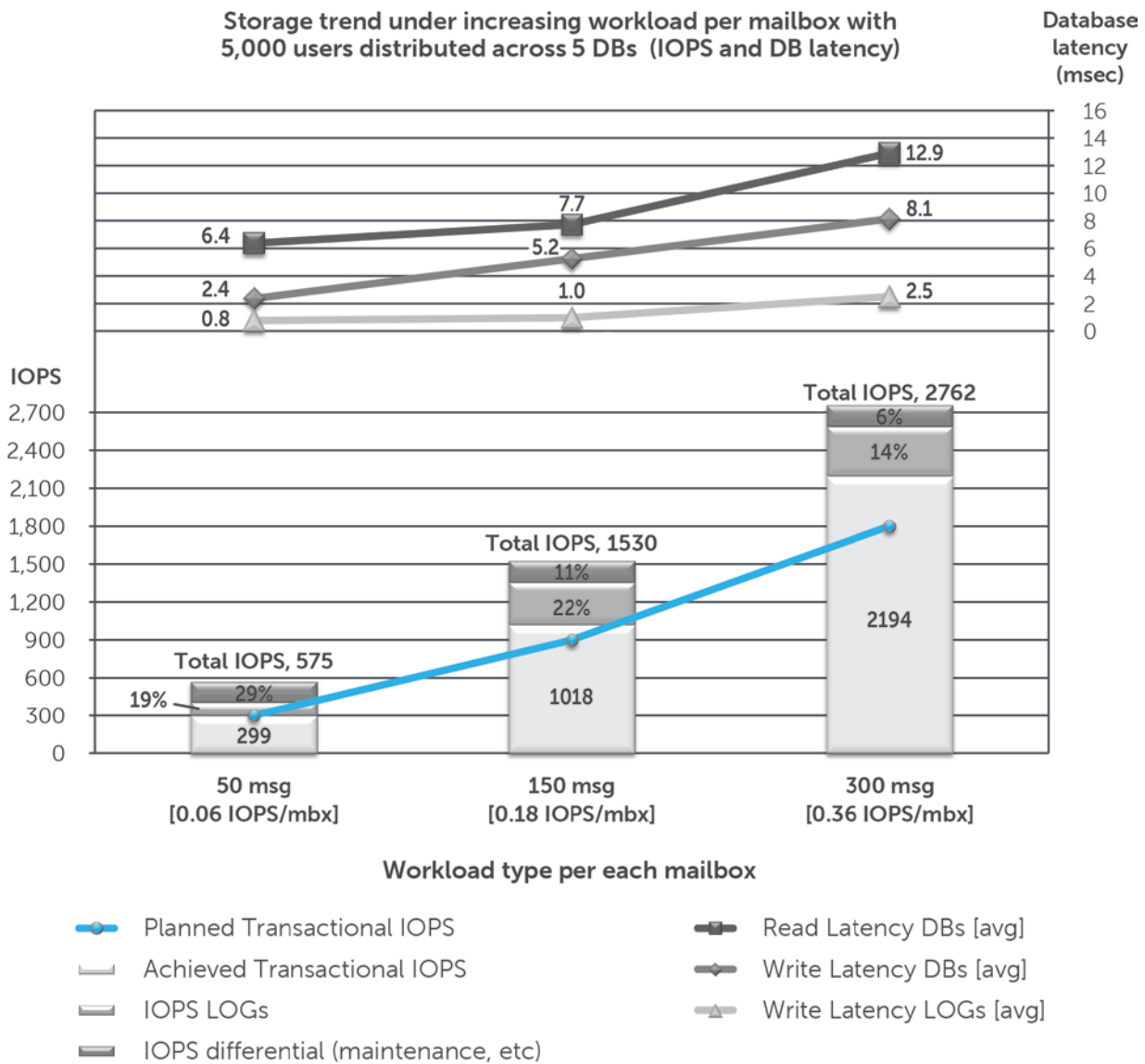
Formerly, the metric adopted for Exchange user activities was founded on light to heavy classification with intermediate degrees of workload. Currently, the user profiles are identified by the association of the number of messages sent and received per mailbox per day, where the average message size is 75KB, and the corresponding amount of memory allocated in the database cache per each mailbox.

The three profiles selected for the test, reported in the Table 2, specify an average mailbox workload progressing from 50 messages to 150 and then to 300 messages per day. The corresponding average cache allocations for the same mailbox are estimated to be 3 MB, 6 MB, and 18 MB. The final estimated IOPS per mailbox, according to these profiles is evaluated in 0.06, 0.18, and 0.36 IOPS.

When applying this count to our simulated scenario of 5,000 mailboxes, the industry accepted average user count for a medium enterprise, hosted by a single server, we account a total estimated IOPS progression (transactional only) of 300, 900, and 1,800 IOPS. These estimated IOPS values do not rate all the remaining Exchange Server activities generating IO access. The database cache evaluated for this pool of users would be 15 GB, 30 GB, and 90 GB; again without considering the additional memory requirements due to other factors.

The results collected from the Exchange Jetstress simulation of these three workload profiles are shown in Figure 4.





**Figure 4** Storage trend under increasing workload per mailbox with 5,000 users distributed across 5 DBs

**Note:** When simulating a different workload with Exchange Jetstress, the tool increases and decreases the generated IOPS via a couple of tuning parameters (threads and sluggish sessions). It is not always viable to exactly match the planned with the achieved IOPS because sometimes the number of threads used distributes more requests. This happens regardless of the effort to configure Exchange Jetstress to perform the exact IOPS planned. The resultant set of latency values must be regarded as the performance obtained by achieved and not planned IOPS. If we evaluate the latency values as an absolute result, they should be equated from achieved versus planned IOPS through a percentage based equation.

During the workload analysis we observed some behavioral changes worth mentioning:

- The reads/writes ratio changes with the workload. The heavier the load, the more writes per seconds were recorded; with an even average of 50% reads /50% writes while running the reference workload profile of 150 messages per day, a 58%/42% ratio for the 50 messages profile and 48%/52% ratio for the 300 messages profile.

- The IOPS differential represented in Figure 4 is almost entirely due to the database maintenance activities. The amount of IOPS generated to perform the maintenance for our database was constant. When evaluated as a percentage, the maintenance load had a heavy 29% impact in the lower workload scenario (0.06 IOPS per mailbox), but decreased to 11% and then 6% in the remaining simulations demanding more transactional IOPS to complete.
- The latency values recorded were well below the Microsoft advised threshold of 20 msec for the DB reads and writes and 10 msec for the log writes even under the heaviest load of 300 messages per day profile.

Table 3 shows an example of converting the resultant set of latency values to match the discrepancy between planned versus achieved IOPS.

**Table 3 Test results: example of latency value conversions to match planned versus achieved IOPS**

|                            | <b>50 msg / 0.06 IOPS</b> | <b>150 msg / 0.18 IOPS</b> | <b>300 msg / 0.36 IOPS</b> |
|----------------------------|---------------------------|----------------------------|----------------------------|
| IOPS achieved <=> planned  | 299 IOPS [100%]           | 1018 IOPS [113%]           | 2194 IOPS [122%]           |
| Read Latency DBs [avg]     | 6.4 msec                  | 7.7 msec                   | 12.9 msec                  |
| Write Latency DBs [avg]    | 2.4 msec                  | 5.2 msec                   | 8.1 msec                   |
| Write latency LOGs [avg]   | 0.8 msec                  | 1.0 msec                   | 2.5 msec                   |
| IOPS equated = planned     | 300 IOPS [100%]           | 900 IOPS [100%]            | 1800 IOPS [100%]           |
| Read Latency DBs equated   | unchanged                 | 6.8 msec                   | 10.6 msec                  |
| Write Latency DBs equated  | unchanged                 | 4.6 msec                   | 6.7 msec                   |
| Write latency LOGs equated | unchanged                 | 0.9 msec                   | 2.1 msec                   |

Table 4 lists the relative improvement or decline of performance recorded during the increase of workload. The percentages are calculated against the total IOPS shown in Figure 4, not just the transactional IOPS. The relevant metrics to evaluate Exchange responsiveness (database or log latency) must be derived from the entire amount of access performed against the storage subsystem, as the transactional load is merely a subset of the total.

**Table 4 Test results: improvement or decline relationship of performance under workload increase**

|                      | <b>50 msg / 0.06 IOPS</b> | <b>150 msg / 0.18 IOPS</b> | <b>300 msg / 0.36 IOPS</b> |
|----------------------|---------------------------|----------------------------|----------------------------|
| IOPS planned         | 100% [300 IOPS]           | 300% [900 IOPS]            | 600% [1800 IOPS]           |
| Total IOPS performed | 100% [575 IOPS]           | 266% [1530 IOPS]           | 480% [2762 IOPS]           |
| Read Latency DBs     | 100% [6.4 msec]           | 121% [7.7 msec]            | 203% [12.9 msec]           |
| Write Latency DBs    | 100% [2.4 msec]           | 223% [5.2 msec]            | 345% [8.1 msec]            |
| Write latency LOGs   | 100% [0.8 msec]           | 125% [1.0 msec]            | 323% [2.5 msec]            |

The outcomes exhibit a powerful trend where the Exchange storage performance indicators (latencies) report a performance penalty consistently lower than the increase of workload. Multiplying by three or six times the workload did not equate in a similar increment in latency. Furthermore, the write latencies, both database and logs, grew more than the read latencies, which confirms the trend of the read/write ratio previously underlined (while write percentage increased with the overall

workload). The dispensed IOPS, and resultant latencies, for the second and third tests would show more inferior values if they were equated following the argument exposed and represented in Table 3.

## 5.2 Characterize the mailbox size

The goal of the mailbox size variance analysis was to establish the storage trend, IOPS ratios, and relationship while varying the size of the user mailboxes and maintaining the remaining factors. The configuration parameters for the test are shown in Table 5.

**Table 5 Test parameters: varying mailbox size**

| Reference configuration: variable factor        |   |
|---|---|
| Mailbox size                                    | 500 MB / 1 GB / 1.5 GB per mailbox          |
| Databases size                                  | 500 GB / 1 TB / 1.5 TB each                 |
| Reference configuration: unchanged factors      |   |
| Number of simulated mailboxes/users             | 5,000 concurrent user                       |
| Number of databases                             | 5 databases (active)                        |
| Mailbox allocation                              | 1,000 mailboxes per each mailbox database   |
| Number of database replica copies               | 1 (standalone)                              |
| IOPS per mailbox / messages per day per mailbox | 0.18 IOPS / 150 messages                    |
| RAID policy                                     | RAID 50                                     |
| Array model, amount of units, SAN configuration | 1x PS6100XV 3.5", one single pool (default) |

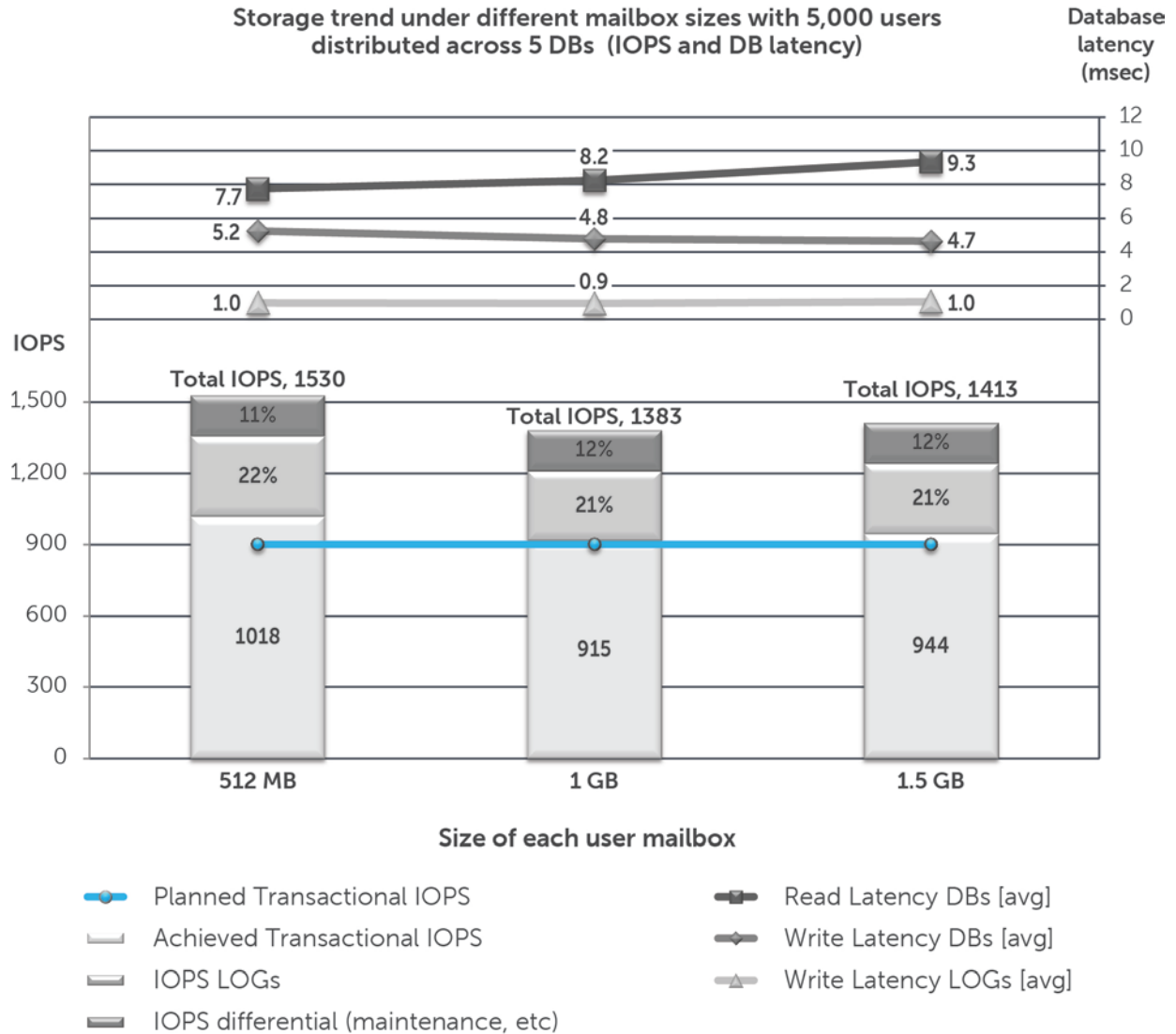
Exchange Server databases size is determined by multiple factors. The main component affecting the capacity requirements is the average mailbox size per user, known as mailbox storage quota, where the term *quota* is usually correlated with the proposition of enforcing a maximum space or capacity limit. The estimate to plan the database file size is easily completed by multiplying the mailbox storage quota by the number of mailboxes per database (e.g. 700 mailboxes with a mailbox storage quota of 800 MB equate to a minimum database capacity requirement of 5.6 TB).

The supplementary factors concurring in the database capacity requirements are the dumpster size, directly related to each mailbox, and the database empty space (an offshoot of mailbox usage and database maintenance activity). The mailbox dumpster contains the items deleted from the end users, but not yet purged from the Exchange system according to a set of retention rules (retention window per mailbox items, per calendar items, single item recovery). As the maximum size of the database is not set, the continuous user activity allocates database pages where free or eventually expands the database file if new ones are needed. While the database maintenance frees tombstoned objects and the online defragmentation consolidates the user data optimizing the B-tree, the empty space in the database is assembled.

For additional information about Exchange mailbox capacity factors refer to Microsoft documentation: *Understanding Mailbox Database and Log Capacity Factors*, available at: <http://technet.microsoft.com/en-us/library/ee832796.aspx>

The three mailbox storage quotas selected for the test (and listed in the Table 5) specify an average mailbox size expanded from 512 MB to 1 GB, and then 1.5 GB. The resultant database file size for our scenario of 1,000 users per database becomes respectively 512 GB, 1 TB, and 1.5 TB.

The results collected from the Exchange Jetstress simulation of these three simulations are reported in Figure 5.



**Figure 5 Storage trend under different mailbox sizes with 5,000 users distributed across 5 DBs**

Table 6 lists the relative improvement or decline of performance recorded during the growth of mailbox size. The percentages are calculated against the total IOPS (not just transactional IOPS) shown in Figure 5.

**Table 6 Test results: improvement or decline relationship of performance when growing the mailbox size**

|                      | <b>512MB [500GB DB]</b> | <b>1GB [1TB DB]</b> | <b>1.5GB [1.5TB]</b> |
|----------------------|-------------------------|---------------------|----------------------|
| IOPS planned         | 100% [900 IOPS]         | 100% [900 IOPS]     | 100% [900 IOPS]      |
| Total IOPS performed | 100% [1530 IOPS]        | 90% [1383 IOPS]     | 92% [1413 IOPS]      |
| Read Latency DBs     | 100% [7.7 msec]         | 106% [8.2 msec]     | 121% [9.3 msec]      |
| Write Latency DBs    | 100% [5.2 msec]         | 91% [4.8 msec]      | 89% [4.7 msec]       |
| Write latency LOGs   | 100% [1.0 msec]         | 97% [0.9 msec]      | 107% [1.0 msec]      |

**Note:** For graphing purposes, the numerical results presented in this paper have been rounded as well as zero decimal digits for percentages and IOPS, and one digit for latencies used.

The outcomes communicate a steady behavior for write latencies closely following the trend of IOPS dispensed. The read latencies show instead a mild, but confirmed growth while running in a scenario with larger mailboxes. The trend is largely expected considering that in the second and third scenario the disk surface to be reached by the drive spindle movements was doubled and then tripled. The slight decrease in achieved IOPS is mostly due to the load generator tool, which was tuned to provide the closest amount of workload, but performed differently in each test because of the wider amount of data to be accessed by the working threads.

### 5.3 Study the database volumes layout

The goal of the database deployment layout analysis was to establish the storage trend, IOPS ratios, and relationship while changing the number of mailbox databases (keeping SAN volume/database ratio 1:1) and maintaining the remaining factors. The configuration parameters for the test are shown in Table 7.

**Table 7 Test parameters: changing databases layout**

| <b>Reference configuration: variable factor</b>   |   |
|---|---|
| Number of databases                               | 2 / 5 / 10 /20 databases (active)                             |
| Databases size                                    | 1.25 TB / 500 GB / 250 GB / 125 GB each                       |
| Mailbox allocation                                | 2,500 / 1,000 / 500 / 250 mailboxes per each mailbox database |
| <b>Reference configuration: unchanged factors</b> |   |
| Number of simulated mailboxes/users               | 5,000 concurrent users  |
| Mailbox size                                      | 512 MB  |
| Number of database replica copies                 | 1 (standalone)  |
| IOPS per mailbox / messages per day per mailbox   | 0.18 IOPS / 150 messages                                      |
| RAID policy                                       | RAID 50   |
| Array model, amount of units, SAN configuration   | 1x PS6100XV 3.5", one single pool (default)                   |

Exchange Server provides the ability to implement multiple databases on the same server with a limit imposed by the licensed version of Exchange Server installed (a maximum of 5 databases for Exchange Standard edition and a maximum of 100 databases for Exchange Enterprise edition). The number of

user mailboxes hosted in a single database is not bound by a declared limit. It is an informed decision of the Exchange administrators to select how many users per database are deployed. The online mailbox move feature allows administrators to seamlessly redistribute mailboxes across databases or mailbox servers as well as postpone or lift constraints imposed by decisions made early in the design and deployment stage of the messaging infrastructure.

The four deployment scenarios selected for the test (and listed in the Table 7) specify an increasing number of databases while maintaining the total amount of users per mailbox server at 5,000. The progressive changes in the configurations include number of databases, users per database, and resultant database size shifting from two databases (2,500 users/DB, 1.25TB DB), five databases (1,000 users/DB, 500GB DB), ten databases (500 users/DB, 250GB DB), and then to 20 databases (250 users/DB, 125GB DB).

The results collected from the Exchange Jetstress simulation of these four workload profiles are displayed in Figure 6.

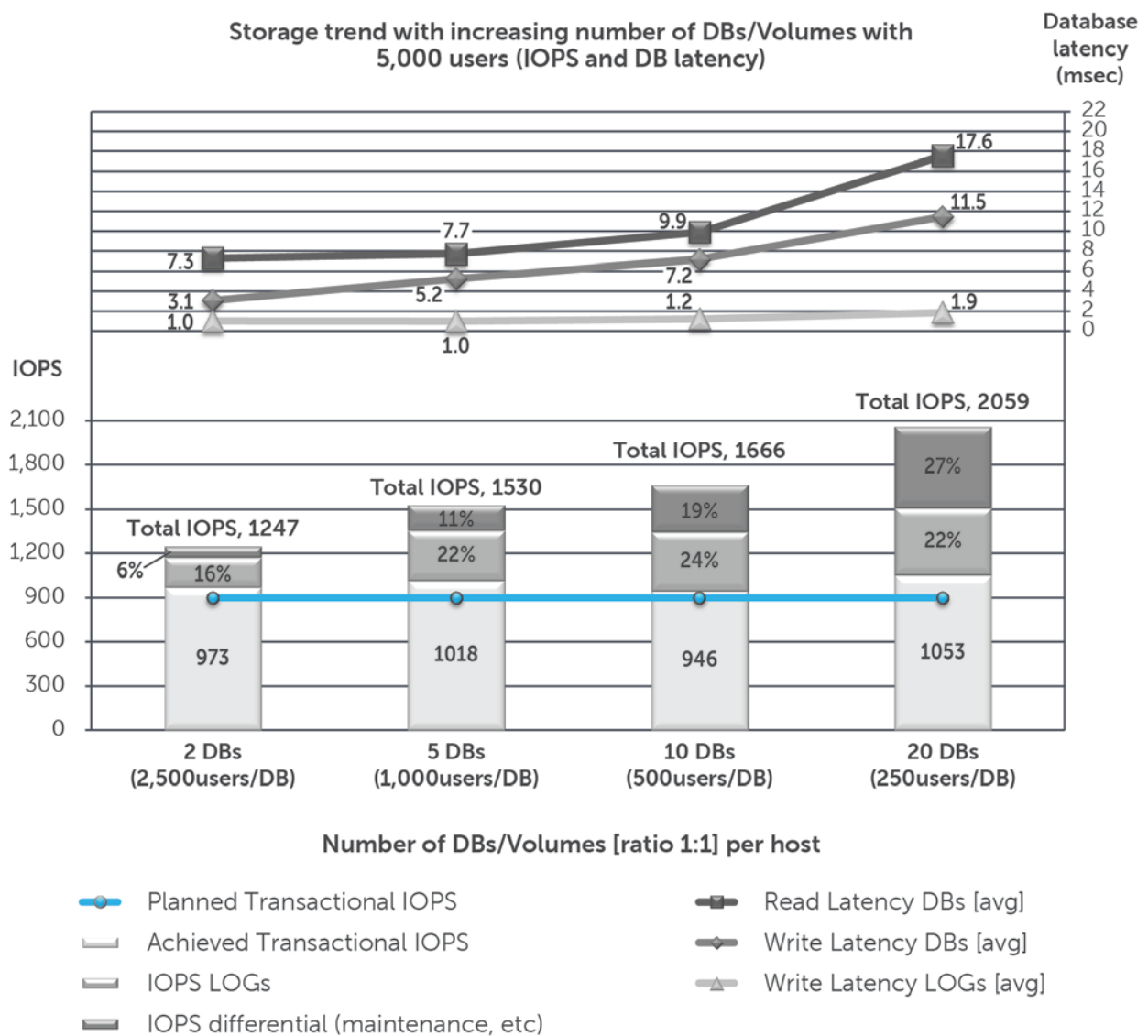


Figure 6 Storage trend with increasing number of DBs/Volumes with 5,000 users

Table 8 lists the relative improvement or decline of performance recorded when increasing the number of databases. The percentages are calculated against the total IOPS (not just the transactional IOPS) shown in Figure 6.

**Table 8 Test results: improvement or decline relationship of performance when increasing the DB number**

|                      | <b>2 DBs</b>     | <b>5 DBs</b>     | <b>10 DBs</b>    | <b>20DBs</b>     |
|----------------------|------------------|------------------|------------------|------------------|
| IOPS planned         | 100% [900 IOPS]  | 100% [900 IOPS]  | 100% [900 IOPS]  | 100% [900 IOPS]  |
| Total IOPS performed | 100% [1247 IOPS] | 123% [1530 IOPS] | 134% [1666 IOPS] | 165% [2059 IOPS] |
| Read Latency DBs     | 100% [7.3 msec]  | 106% [7.7 msec]  | 136% [9.9 msec]  | 242% [17.6 msec] |
| Write Latency DBs    | 100% [3.1 msec]  | 171% [5.2 msec]  | 234% [7.2 msec]  | 373% [11.5 msec] |
| Write latency LOGs   | 100% [1.0 msec]  | 98% [1.0 msec]   | 121% [1.2 msec]  | 186% [1.9 msec]  |

By default, Exchange Server 2010 activates the maintenance tasks on each new mailbox database created. Most of these activities concur to the health and performance of the databases and would not be deactivated. Only the online blocks checksum can be optionally offloaded in a scheduled time window if required. The maintenance overhead must be considered when increasing the number of databases, while a careful consideration is given when selecting large databases offering a lower granularity of administration and data protection.

The outcomes expose a uniform trend where the number of databases implemented considerably influences the total IOPS required first, and then the overall performance of the deployment. All three latency indicators report a significant growth up to ten databases, and then further amplify the performance decline in the test using 20 databases. For instance, an additional test with 40 databases was attempted, but did not pass the previously defined latency criteria. The database maintenance impact is the measurable key element affecting the amount of IOPS dispensed, routinely widening the percentage of IOPS from 6% to 27% over a quarter of all the IOPS in the fourth test.

## 5.4 Study the DAG number of database copies differential

The goal of the high availability scenarios test was to establish the storage trend, and IOPS ratios, and relationship while manipulating the DAG technology through multiple database replica copies and maintaining the remaining factors. The configuration parameters for the test are shown in Table 9.

**Table 9 Test parameters: increasing number of database replica copies**

| Reference configuration: variable factor        |   |
|---|---|
| Number of database replica copies               | 1 (standalone) / 2 nodes / 4 nodes / 8 nodes<br>Database Availability Group |
| Reference configuration: unchanged factors      |   |
| Number of simulated mailboxes/users             | 5,000 concurrent users  |
| Mailbox size                                    | 512 MB  |
| Number of databases                             | 5 databases (active)  |
| Databases size                                  | 500 GB each   |
| Mailbox allocation                              | 1,000 mailboxes per each mailbox database                                   |
| IOPS per mailbox / messages per day per mailbox | 0.18 IOPS / 150 messages  |
| RAID policy                                     | RAID 50   |
| Array model, amount of units, SAN configuration | 1x PS6100XV 3.5", one single pool (default)                                 |

Exchange Server 2010 introduces the DAG technology to mitigate the risks of service downtime or data loss due to the corruption of mailbox databases or failure of mailbox servers. The DAG enables one or more databases to be replicated across several servers, with no data loss between the nodes of the cluster group, while only one copy of the database(s) is active and accessible from the end users at a specific point in time. Exchange Replication service manages the activities inherent to the replica process after a log file is closed on the node hosting the active or source database and the respective log replay process on the node hosting a passive or target database copy. Regardless of whether the databases are in an active status or not, the maintenance tasks run generating IOPS toward the storage subsystem.

A user mailbox hosted on a server and protected by a DAG configuration generates less transactional IOPS (minus 20% on average) than the equivalent in a standalone environment. Most of the difference is caused by the change in behavior of the Exchange database cache and log checkpoint depth as discussed in section 3.2. The 0.18 IOPS per mailbox workload was not adjusted and lowered during the DAG scenarios test subset because we preferred to maintain a numerical output comparable with the rest of the test sets.

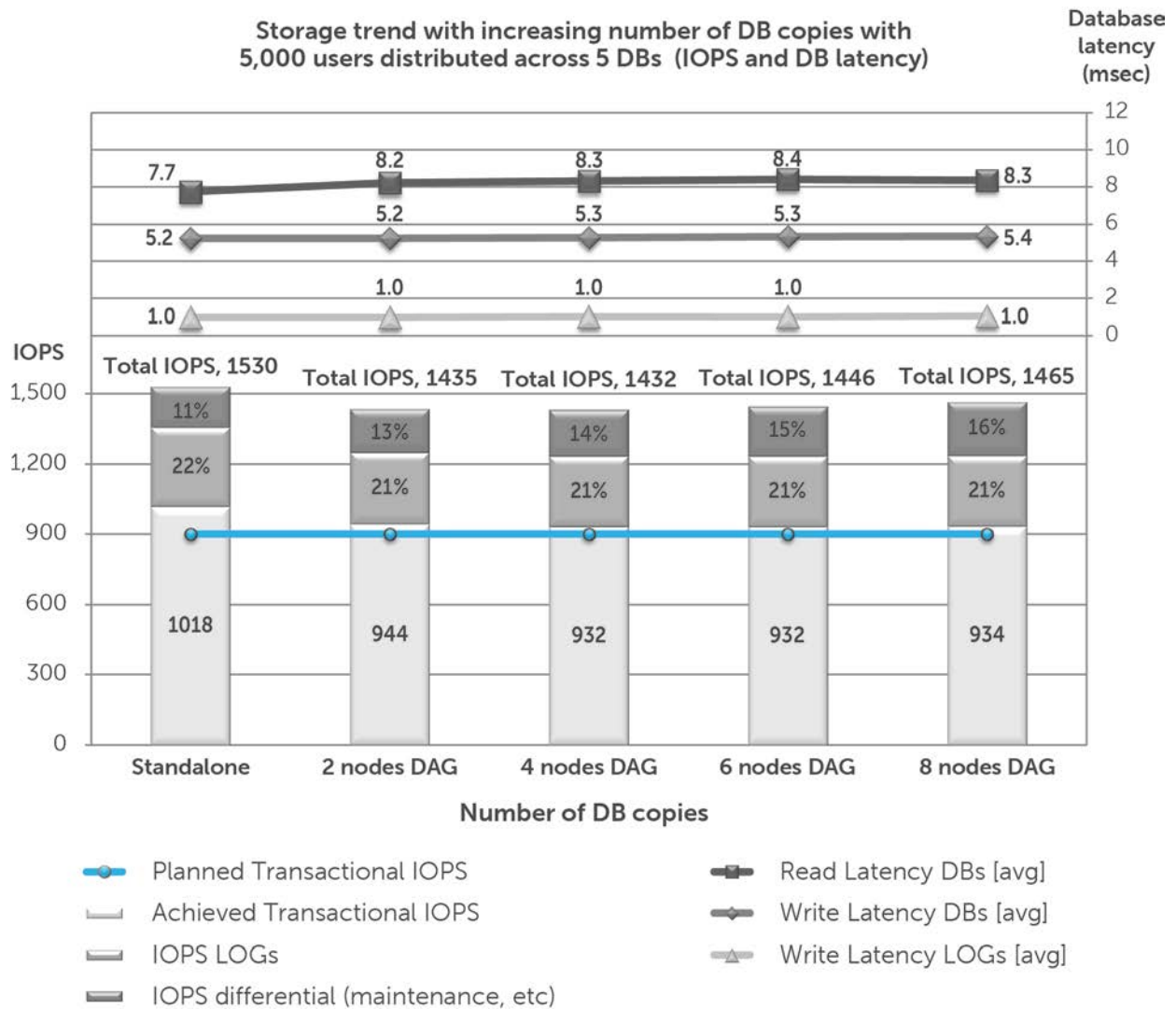
Although Exchange DAG supports two types of replica copies (highly available database copies and lagged database copies), the Exchange Jetstress simulation tool only reproduces the storage access produced against the active copy of a replicated database.

A study of every aspect and dependency for a resilience model of the mailbox server role is beyond the scope of this paper. Nevertheless, the five scenarios from standalone server to multiple nodes DAG, reported in Table 9, identify a broad set of deployment cases with database replica copies: standalone server (with five active database copies), two nodes DAG (with five active database copies and five replicated), four nodes DAG (with five active database copies and 15 replicated), six nodes



DAG (with five active database copies and 25 replicated), and eight nodes DAG (with five active database copies and 35 replicated).

The results collected from the Exchange Jetstress simulation of these five scenarios are reported in Figure 7.



**Figure 7 Storage trend with increasing number of DBs copies with 5,000 users distributed across 5 DBs**

Table 10 reports the relative improvement or decline of performance recorded with an incremental number of database replica copies. The percentages are calculated against the total IOPS shown in Figure 7, not only the transactional IOPS.

**Table 10 Test results: improvement or decline relationship of performance with an incremental number of database replica copies**

|                      | Standalone          | 2 nodes            | 4 nodes            | 6 nodes            | 8 nodes            |
|----------------------|---------------------|--------------------|--------------------|--------------------|--------------------|
| IOPS planned         | 100%<br>[900 IOPS]  | 100%<br>[900 IOPS] | 100%<br>[900 IOPS] | 100%<br>[900 IOPS] | 100%<br>[900 IOPS] |
| Total IOPS performed | 100%<br>[1530 IOPS] | 94%<br>[1435 IOPS] | 94%<br>[1432 IOPS] | 95%<br>[1446 IOPS] | 96%<br>[1465 IOPS] |
| Read Latency DBs     | 100%<br>[7.7 msec]  | 106%<br>[8.2 msec] | 107%<br>[8.3 msec] | 108%<br>[8.4 msec] | 108%<br>[8.3 msec] |
| Write Latency DBs    | 100%<br>[5.2 msec]  | 100%<br>[5.2 msec] | 101%<br>[5.3 msec] | 101%<br>[5.3 msec] | 102%<br>[5.4 msec] |
| Write latency LOGs   | 100%<br>[1.0 msec]  | 100%<br>[1.0 msec] | 104%<br>[1.0 msec] | 103%<br>[1.0 msec] | 107%<br>[1.0 msec] |

The outcomes display a steady trend where the only perceivable pattern when incrementing the number of database copies was a slight increase in IOPS spent outside the transactional and logs writing activities (reported as IOPS differential in Figure 7). The supplementary reads processed against the logs to simulate the Exchange Replication service logs shipping method were responsible for this increase, simultaneously reflected in the minor latency rise. Furthermore, with the addition of database replica copies we recorded a minimal variation in the reads/writes ratio from the original 50% reads / 50% writes up to 52% reads / 48% writes.

## 5.5 Characterize the RAID policy

The goal of this test was to establish the storage trend, and IOPS ratios and relationship while implementing different RAID policies on the EqualLogic SAN array and maintaining unchanged the remaining factors. The configuration parameters for the test are shown in Table 11.

**Table 11 Test parameters: RAID policies variation**

| Reference configuration: variable factor        |   |
|---|---|
| RAID policy                                     | RAID 10 / RAID 50 / RAID 5 / RAID 6         |
| Reference configuration: unchanged factors      |   |
| Number of simulated mailboxes/users             | 5,000 concurrent users                      |
| Mailbox size                                    | 512MB                                       |
| Number of databases                             | 5 databases (active)                        |
| Databases size                                  | 500GB each                                  |
| Mailbox allocation                              | 1,000 mailboxes per each mailbox database   |
| Number of database replica copies               | 1 (standalone)                              |
| IOPS per mailbox / messages per day per mailbox | 0.18 IOPS / 150 messages                    |
| Array model, amount of units, SAN configuration | 1x PS6100XV 3.5", one single pool (default) |

The EqualLogic PS Series (6100 and 4100) arrays provide a range of different RAID policies and the chance to exploit spare drives to protect data. Each RAID level offers a distinct set of performance and availability characteristics dependent from the nature of the RAID policy and the workload applied.

- **RAID 10**, or striping over multiple mirrored sets (RAID 1)  
The best performance for random access at the expense of an average capacity available.  
A minimal impact for RAID reconstruction in case of drive failure.  
A total of 22 available drives on PS6100, 10 available drives on PS4100, with two spare drives.
- **RAID 50**, or striping over multiple distributed parity sets (RAID 5)  
A balance between performance and capacity.  
A moderate impact for RAID reconstruction in case of drive failure.  
A total of 22 available drives on PS6100, 10 available drives on PS4100, with two spare drives.
- **RAID 5**, one or more distributed parity sets  
The best capacity with some drawbacks in performance for random writes.  
A moderate/heavy impact for RAID reconstruction in case of drive failure.  
A total of 23 available drives on PS6100, 11 available drives on PS4100, with one spare drive.
- **RAID 6**, one or more dual parity sets  
The best availability at the expense of random writes performance and rebuild impact.  
A heavy impact for RAID reconstruction in case of drive failure.  
A total of 23 available drives on PS6100, 11 available drives on PS4100, with one spare drive.

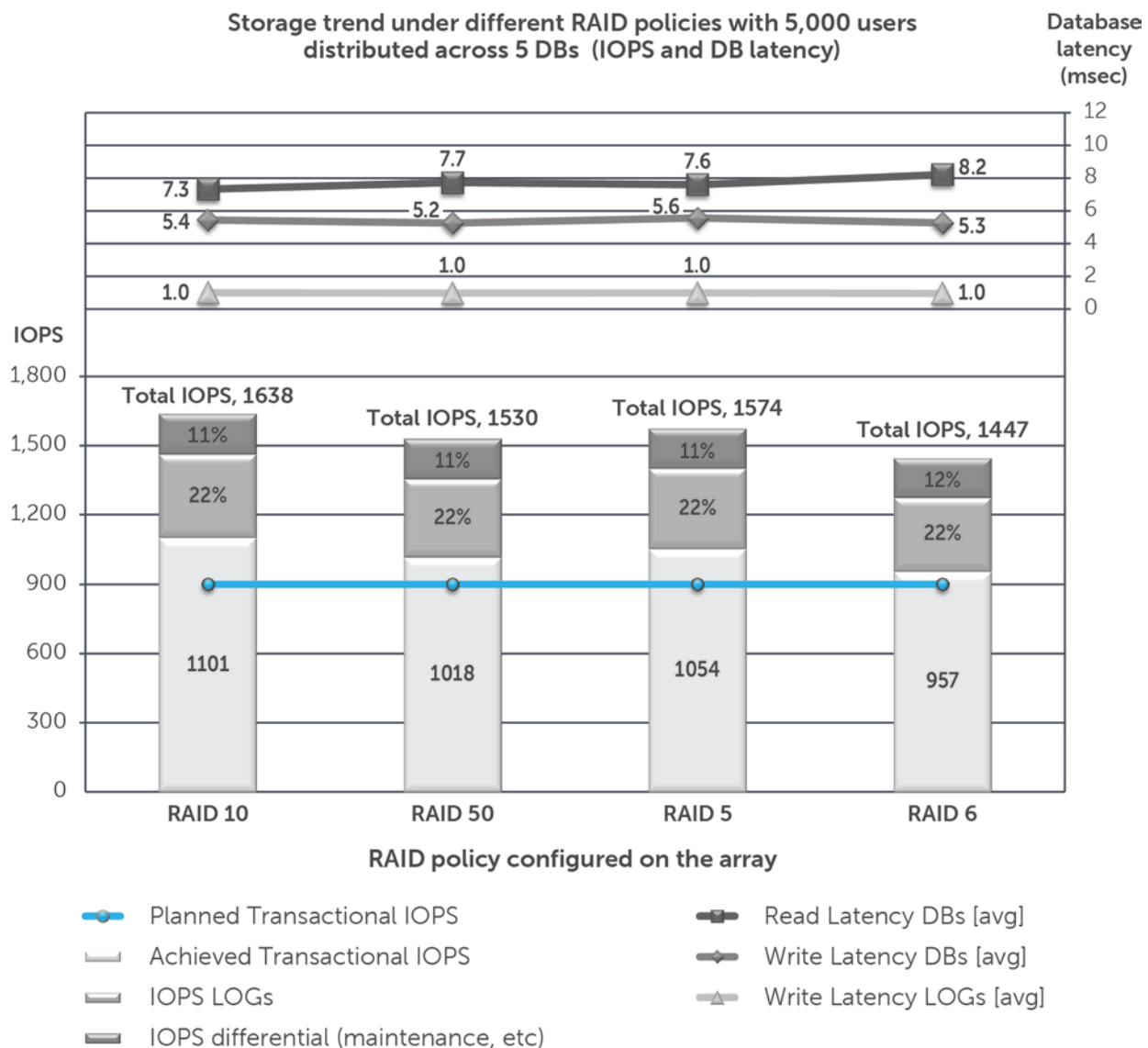
**Note:** RAID implementations deprived of spare drives for protection, even if allowed on the EqualLogic PS Series, are not considered here and in the rest of this study.

EqualLogic PS arrays also support RAID policy conversion allowing the switchover from one RAID level to another without disruption of data presents on the volumes, although the performance during the conversion may be degraded. The following rules apply for RAID conversion:

- Conversion from RAID 10 to any other RAID level
- Conversion from RAID 50 to RAID 5 or 6 levels
- Conversion from RAID 6 to RAID 5 level
- Conversion from RAID 5 is not supported

The four RAID policies selected for the test, reported in the Table 11, use the PS6100XV 3.5" as the reference array. The array was reinitialized to the factory level for each RAID level run, the volumes recreated and the Exchange Jetstress databases redeployed. Specifically for this set of tests the configuration of Exchange Jetstress, in term of threads and sluggish sessions, remains strictly the same across the runs because the aim is to rather capture the IOPS dispensed by the array consistently with the Exchange latency indicators.

The results collected from the Exchange Jetstress simulation against the different RAID levels are reported in Figure 8.



**Figure 8 Storage trend under RAID policies with 5,000 users distributed across 5 DBs**

Table 12 reports the relative improvement or decline of performance recorded while moving to a different RAID implementation on the EqualLogic array. The percentages are calculated against the total IOPS shown in Figure 8, not only the transactional IOPS.

**Table 12 Test results: improvement or decline relationship of performance under different RAID policies**

|                      | RAID 10          | RAID 50         | RAID 5          | RAID 6          |
|----------------------|------------------|-----------------|-----------------|-----------------|
| IOPS planned         | 100% [900 IOps]  | 100% [900 IOps] | 100% [900 IOps] | 100% [900 IOps] |
| Total IOPS performed | 100% [1638 IOps] | 93% [1530 IOps] | 96% [1574 IOps] | 88% [1447 IOps] |
| Read Latency DBs     | 100% [7.3 msec]  | 106% [7.7 msec] | 104% [7.6 msec] | 112% [8.2 msec] |
| Write Latency DBs    | 100% [5.4 msec]  | 97% [5.2 msec]  | 102% [5.6 msec] | 97% [5.3 msec]  |
| Write latency LOGs   | 100% [1.0 msec]  | 97% [1.0 msec]  | 98% [1.0 msec]  | 95% [1.0 msec]  |

The outcomes confirm the superior IOPS performance of RAID 10 level across the entire pool of policies, with a percentage gap of the competitors from 4% of RAID 5 to a maximum of 12% of RAID 6, with RAID 50 at 7% close to RAID 5. The achievements are influenced by the advantage to have a greater number of working spindles for RAID 5/6 (23 drives) when compared to RAID 10/50 (22 drives). The latencies recorded show a regular trend with the two limited deviations of improved RAID 10 and declined RAID 6. In the case the latency results are equated with the methodology conversion shown in section 5.1 the numerical values will be even more pronounced in the direction just remarked.

## 5.6 Scaling up the number of users

The goal of this test was to establish the storage trend, and IOPS ratios and relationship while increasing the number of concurrent users and maintaining unchanged the remaining factors. The configuration parameters for the test are shown in Table 13.

**Table 13 Test parameters: concurrent users scalability**

| Series  | Reference configuration: variable factor    |  |
|---|---|--|
| #1  | Number of simulated mailboxes/users         | 5,000 / 7,500 / 10,000 / 12,500 concurrent users |
|   | Number of databases                         | 2 / 3 / 4 / 5 databases (active)                 |
|   | Databases size                              | 1.5TB each                                       |
|   | Mailbox allocation                          | 2,500 mailboxes per each mailbox DB              |
| #2  | Number of simulated mailboxes/users         | 5,000 / 7,000 / 9,000 / 10,000 concurrent users  |
|   | Number of databases                         | 5 / 7 / 9 / 10 databases (active)                |
|   | Databases size                              | 500GB each                                       |
|   | Mailbox allocation                          | 1,000 mailboxes per each mailbox DB              |
| <b>Reference configuration: unchanged factors</b> |   |  |
| Mailbox size                                      | 512MB                                       |  |
| Number of database replica copies                 | 1 (standalone)                              |  |
| IOPS per mailbox / messages per day per mailbox   | 0.18 IOPS / 150 messages                    |  |
| RAID policy                                       | RAID 50                                     |  |
| Array model, amount of units, SAN configuration   | 1x PS6100XV 3.5", one single pool (default) |  |

The building block used to scale up the number of users for our tests was the addition of databases, while maintaining the users per database ratio stable. The practice of gradually provisioning predefined sized mailbox databases simplifies the administrative burden when the demand to support more users increases and conveniently increments the workload in a linear fashion.

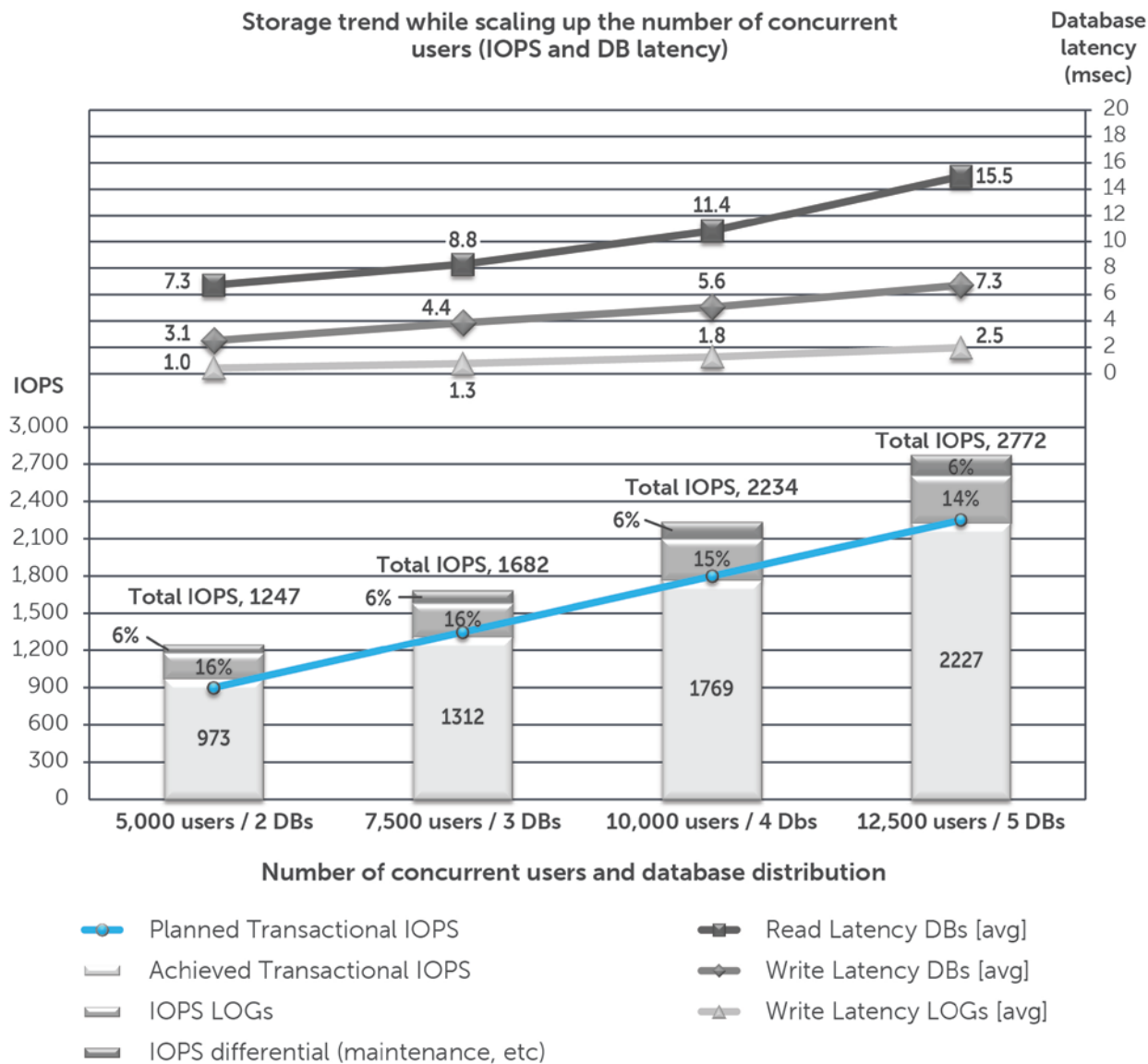
The selection of the mailbox databases layout follows the outcomes presented in section 5.3. The users per mailbox database ratio selected for our series of tests reflects two plausible deployment scenarios: 2,500 and 1,000 respectively. Each series includes four different iterations increased by a progressive amount of workload. First series scales up the amount of concurrent users from 5,000, to 7,500, 10,000 and then 12,500 by incrementally adding mailbox database units of 1.5TB with 2,500 mailboxes each. Second series instead scales from 5,000 to 7,000, 9,000 and then 10,000 by adding mailbox database units of 500GB containing 1,000 mailboxes each.

The database cache evaluated for the pool of users in each series is reported in Table 14, again without considering the additional memory requirements due to other factors. The estimates for the database cache are based on Microsoft published metrics, and not on recorded values from our tests. Microsoft Exchange Jetstress resources utilization performs differently from an Exchange Server installation for both memory and processing resources, as reported in Appendix B.

**Table 14 Exchange database cache (estimated) while scaling up the number of users**

| Series    | Number of concurrent users | Cache per user | Exchange database Cache (estimated) |
|-----------|----------------------------|----------------|-------------------------------------|
| <b>#1</b> | 5,000                      | 6MB            | 30GB                                |
|           | 7,000                      |                | 42GB                                |
|           | 9,000                      |                | 54GB                                |
|           | 10,000                     |                | 60GB                                |
| <b>#2</b> | 5,000                      | 6MB            | 30GB                                |
|           | 7,500                      |                | 45GB                                |
|           | 10,000                     |                | 60GB                                |
|           | 12,500                     |                | 75GB                                |

The results collected from the Exchange Jetstress simulation of the two series of incremental workload are reported in Figure 9 and Figure 10.



**Figure 9 Storage trend while scaling up the number of concurrent users (2,500 users per database)**

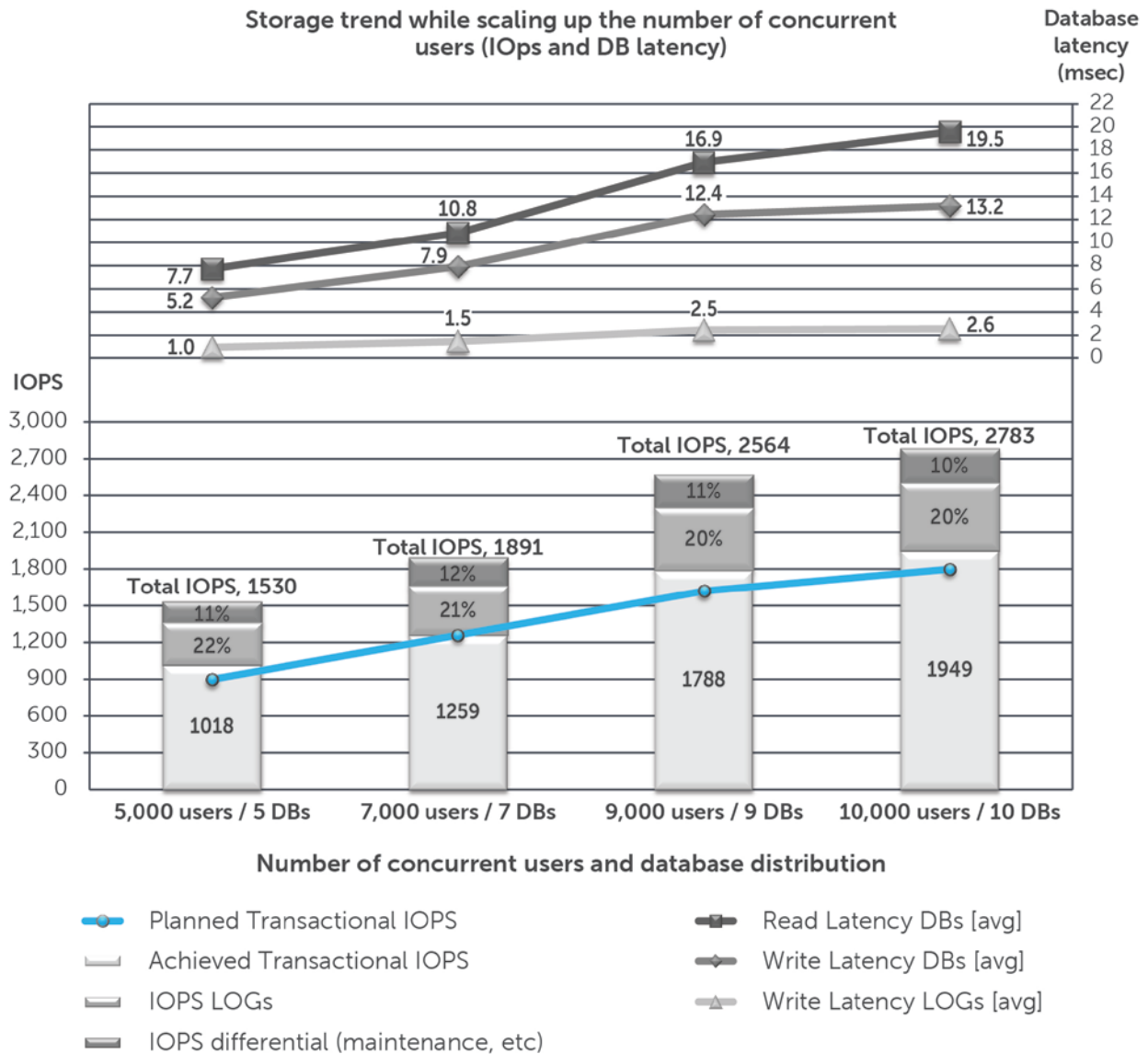
Table 15 reports the relative improvement or decline of performance recorded while incrementally adding concurrent users and mailbox databases during test series #1. The percentages are calculated against the total IOPS shown in Figure 9, not only the transactional IOPS.

**Table 15 Test results: improvement or decline relationship of performance while scaling up concurrent users in test series #1**

| <b>Series #1</b>     | <b>5,000 users</b> | <b>7,500 users</b> | <b>10,000 users</b> | <b>12,500 users</b> |
|----------------------|--------------------|--------------------|---------------------|---------------------|
| IOPS planned         | 100% [900 IOps]    | 150% [1350 IOps]   | 200% [1800 IOps]    | 250% [2250 IOps]    |
| Total IOPS performed | 100% [1247 IOps]   | 135% [1682 IOps]   | 179% [2234 IOps]    | 222% [2772 IOps]    |
| Read Latency DBs     | 100% [7.3 msec]    | 122% [8.8 msec]    | 156% [11.4 msec]    | 214% [15.5 msec]    |
| Write Latency DBs    | 100% [3.1 msec]    | 144% [4.4 msec]    | 183% [5.6 msec]     | 237% [7.3 msec]     |
| Write latency LOGs   | 100% [1.0 msec]    | 134% [1.3 msec]    | 184% [1.8 msec]     | 254% [2.5 msec]     |

The outcomes of the first series reveal a linear scalability retaining the key performance indicators abundantly below the reference Microsoft thresholds, and further show an increase in the recorded latencies constantly underneath the corresponding percentage increase of the workload. The reads versus writes ratio reported for this database layout is confirmed at 54% reads / 46% writes as in the section 5.3 test for the same layout.





**Figure 10 Storage trend while scaling up the number of concurrent users (1,000 users per database)**

Table 16 reports the relative improvement or decline of performance recorded while incrementally adding concurrent users and mailbox databases during test series #2. The percentages are calculated against the total IOPS shown in Figure 10, not only the transactional IOPS.

**Table 16 Test results: improvement or decline relationship of performance while scaling up concurrent users in test series #2**

| Series #2            | 5,000 users      | 7,000 users      | 9,000 users      | 10,000 users     |
|----------------------|------------------|------------------|------------------|------------------|
| IOPS planned         | 100% [900 IOPs]  | 140% [1260 IOPs] | 180% [1620 IOPs] | 200% [1800 IOPs] |
| Total IOPS performed | 100% [1530 IOPs] | 124% [1891 IOPs] | 168% [2564 IOPs] | 182% [2783 IOPs] |
| Read Latency DBs     | 100% [7.7 msec]  | 140% [10.8 msec] | 219% [16.9 msec] | 253% [19.5 msec] |
| Write Latency DBs    | 100% [5.2 msec]  | 151% [7.9 msec]  | 237% [12.4 msec] | 251% [13.2 msec] |
| Write latency LOGs   | 100% [1.0 msec]  | 151% [1.5 msec]  | 252% [2.5 msec]  | 262% [2.6 msec]  |

The outcomes of the second series still expose a linear trend while scaling up the number of concurrent users, and hold the latency metrics beneath the thresholds. The database layout selected, however, affects the absolute performance and the maximum achievable concurrent users per array if compared with the previous series. When the workload moves to 9,000 concurrent users it is perceivable a lightly steep raise in the latencies versus workload percentages, that eases moving forward to the 10,000 concurrent users iteration. The reads versus writes ratio is confirmed at 50% / 50% as in our reference layout and load per mailbox user.

## 5.7 Scaling out the SAN

The goal of this test was to establish the storage trend, and IOPS ratios and relationship while scaling out the SAN by incrementing the number of arrays in order to support an increased number of concurrent users and maintaining unchanged the remaining factors. The configuration parameters for the test are shown in Table 17.

**Table 17 Test parameters: scale up the SAN and the number of users**

| Series  | Reference configuration: variable factor |   |
|---|--|---|
| #1  | Number of simulated mailboxes/users      | 5,000 / 10,000 / 15,000 concurrent users              |
|   | Number of databases                      | 5 / 10 / 15 databases (active)                        |
| #2  | Number of simulated mailboxes/users      | 9,000 / 18,000 / 27,000 concurrent users              |
|   | Number of databases                      | 9 / 18 / 27 databases (active)                        |
| Array model, amount of units, SAN configuration |  | 1x / 2x / 3x PS6100XV 3.5", one single pool (default) |
| Reference configuration: unchanged factors      |  |   |
| Mailbox size                                    |  | 512MB   |
| Databases size                                  |  | 500GB each  |
| Mailbox allocation                              |  | 1,000 mailboxes per each mailbox database             |
| Number of database replica copies               |  | 1 (standalone)  |
| IOPS per mailbox / messages per day per mailbox |  | 0.18 IOPS / 150 messages                              |
| RAID type                                       |  | RAID-50   |

The building block to scale out the SAN to support larger environments is framed around the user per database ratio of 1,000 defined in section 5.1. The three components involved are the EqualLogic SAN array, the Exchange server and the total workload per server, established from the reference workload of 0.18IOPS per mailbox and a set of 5,000 or 9,000 total users per server. The building block ratio of 1:1:1 is conserved across the three steps of the ladder to scale the SAN as summarized in Table 18.

**Table 18 Building block ratio for each test and series regarding scaling out the SAN**

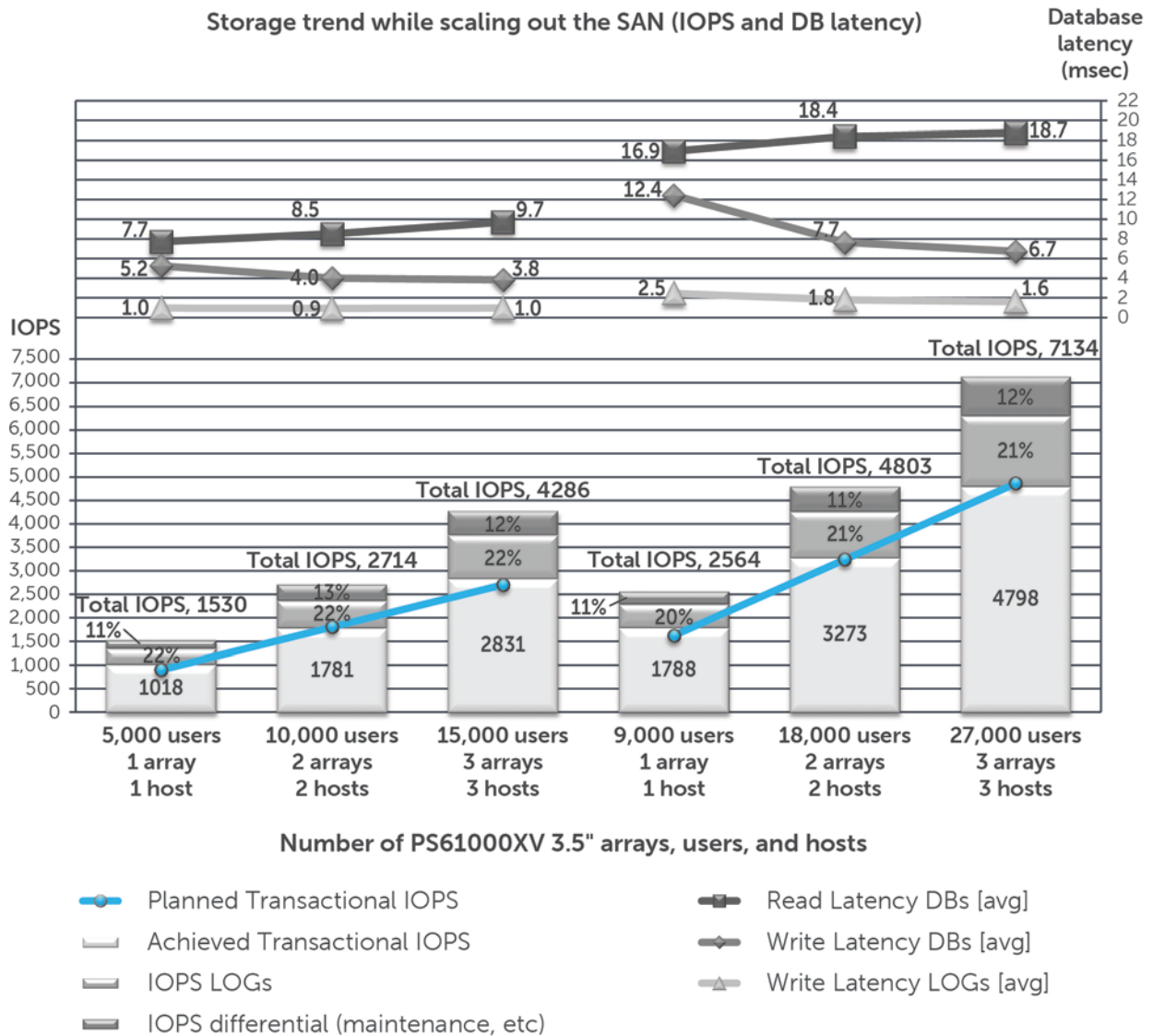
| Series | # of SAN arrays | # of Exchange servers | Workload / database layout       |
|--------|-----------------|-----------------------|----------------------------------|
| #1     | 1               | 1                     | 5,000 concurrent users / 5 DBs   |
| #1     | 2               | 2                     | 10,000 concurrent users / 10 DBs |
| #1     | 3               | 3                     | 15,000 concurrent users / 15 DBs |
| #2     | 1               | 1                     | 9,000 concurrent users / 9 DBs   |
| #2     | 2               | 2                     | 18,000 concurrent users / 18 DBs |
| #2     | 3               | 3                     | 27,000 concurrent users / 27 DBs |

The database cache evaluated for the pool of users in each series and for each Exchange server is reported in Table 19, again without considering the additional memory requirements due to other factors.

**Table 19 Exchange database cache (estimated)**

| Series | Number of concurrent users per host | Cache per user | Exchange database Cache (estimated) per host |
|--------|-------------------------------------|----------------|--|
| #1     | 5,000                               | 6MB            | 30GB   |
| #2     | 9,000                               | 6MB            | 54GB   |

The results collected from the Exchange Jetstress simulation of the two series of incremental workload are reported in Figure 11.



**Figure 11 Storage trend while scaling out the SAN (number of arrays and users)**

Table 20 reports the relative improvement or decline of performance recorded while incrementally adding array units, Exchange servers and concurrent users during test series #1 and #2. The percentages are calculated against the total IOPS shown in Figure 11, not only the transactional IOPS.

**Table 20 Test results: improvement or decline relationship of performance when scaling out the SAN**

| <b>Series #1</b>     | <b>5,000 users<br/>1 array / 1 host</b> | <b>10,000 users<br/>2 arrays / 2 hosts</b> | <b>15,000 users<br/>3 arrays / 3 hosts</b> |
|----------------------|---|--|--|
| IOPS planned         | 100% [900 IOps]                         | 200% [1800 IOps]                           | 300% [2700 IOps]                           |
| Total IOPS performed | 100% [1530 IOps]                        | 177% [2714 IOps]                           | 280% [4286 IOps]                           |
| Read Latency DBs     | 100% [7.7 msec]                         | 110% [8.5 msec]                            | 126% [9.7 msec]                            |
| Write Latency DBs    | 100% [5.2 msec]                         | 76% [4.0 msec]                             | 73% [3.8 msec]                             |
| Write latency LOGs   | 100% [1.0 msec]                         | 96% [0.9 msec]                             | 99% [1.0 msec]                             |
| <b>Series #2</b>     | <b>9,000 users<br/>1 array / 1 host</b> | <b>18,000 users<br/>2 arrays / 2 hosts</b> | <b>27,000 users<br/>3 arrays / 3 hosts</b> |
| IOPS planned         | 100% [1620 IOps]                        | 200% [3240 IOps]                           | 300% [4860 IOps]                           |
| Total IOPS performed | 100% [2564 IOps]                        | 187% [4803 IOps]                           | 278% [7134 IOps]                           |
| Read Latency DBs     | 100% [16.9 msec]                        | 109% [18.4 msec]                           | 111% [18.7 msec]                           |
| Write Latency DBs    | 100% [12.4 msec]                        | 62% [7.7 msec]                             | 54% [6.7 msec]                             |
| Write latency LOGs   | 100% [2.5 msec]                         | 73% [1.8 msec]                             | 66% [1.6 msec]                             |

The outcomes of the both series acknowledge the ability of the EqualLogic SAN to accept and support linearly the additional workload distributed across two then three arrays. While a modest decline in database read latency performance is documented in the trend, the write latencies show a considerable benefit at the addition of PS arrays to the same EqualLogic pool. Notably the two units scenario report a steep and consistent gain in write latency performance, sanctioned more softly by the three arrays scenario.

## 5.8 Assess the 6100 and 4100 family models

The goal of this test was to establish the storage trend and IOPS differential while utilizing different models of EqualLogic PS Series arrays to build the SAN, while keeping factors constant. The configuration parameters for the test are shown in Table 21.

**Table 21 Test parameters: RAID type variation**

| Reference configuration: variable factor        |  |
|---|--|
| Array model, amount of units                    | 1x PS6100XV 3.5"<br>or 1x PS6100X<br>or 1x PS6100E<br>or 1x PS4100XV 3.5"<br>or 1x PS4100E |
| Reference configuration: unchanged factors      |  |
| Number of simulated mailboxes/users             | 5,000 concurrent users (except for PS4100E)  |
| Mailbox size                                    | 512MB  |
| Number of databases                             | 5 databases (active) (except for PS4100E)  |
| Databases size                                  | 500GB each   |
| Mailbox allocation                              | 1,000 mailboxes per each mailbox database  |
| Number of database replica copies               | 1 (standalone)   |
| IOPS per mailbox / messages per day per mailbox | 0.18 IOPS / 150 messages   |
| RAID policy                                     | RAID 50  |
| SAN configuration                               | one single pool (default)  |

The list dispenses the essential details of the EqualLogic PS Series array model (6100 and 4100) used. For additional details refer to the 'Hardware components' list in the Appendix A.

- **PS6100XV 3.5"**: dual controllers with four GbE ports each, and dedicated management port  
Twenty-four 3.5" 15,000RPM SAS disk drives, 600GB each
- **PS6100X**: dual controllers with four GbE ports each, and dedicated management port  
Twenty-four 2.5" 10,000RPM SAS disk drives, 900GB each
- **PS6100E**: dual controllers with four GbE ports each, and dedicated management port  
Twenty-four 3.5" 7,200RPM NL-SAS disk drives, 2TB each
- **PS4100XV 3.5"**: dual controllers with two GbE ports each, and dedicated management port  
Twelve 3.5" 15,000RPM SAS disk drives, 600GB each
- **PS4100E**: dual controllers with two GbE ports each, and dedicated management port  
Twelve 3.5" 7,200RPM NL-SAS disk drives, 2TB each

The main difference between the array models within the PS6100/PS4100 families analyzed during the tests is the drive type. When comparing the two families, the disparity of elements between the two families is significant (i.e. drives count, number of network interfaces per storage controller, maximum scalability). The illustration of the PS4100XV 3.5" within the PS6100 family chart (Figure 12) is a stretched analysis to show the benefits and limitations of the platform. The PS4100 family has then been assessed separately to have a proper validation of the performance within a similar context.

**Note:** The configuration for the physical host and the virtual machine hosting Exchange Jetstress when the PS4100 family was assessed had the number of network adapters (and port group) reduced from four to two. The decrease is matured with the aim of maintaining the 1:1 ratio of network adapters between the array and the host.

The results collected from the Exchange Jetstress simulation against the different PS6100 family models and PS4100XV 3.5" are reported in Figure 12.

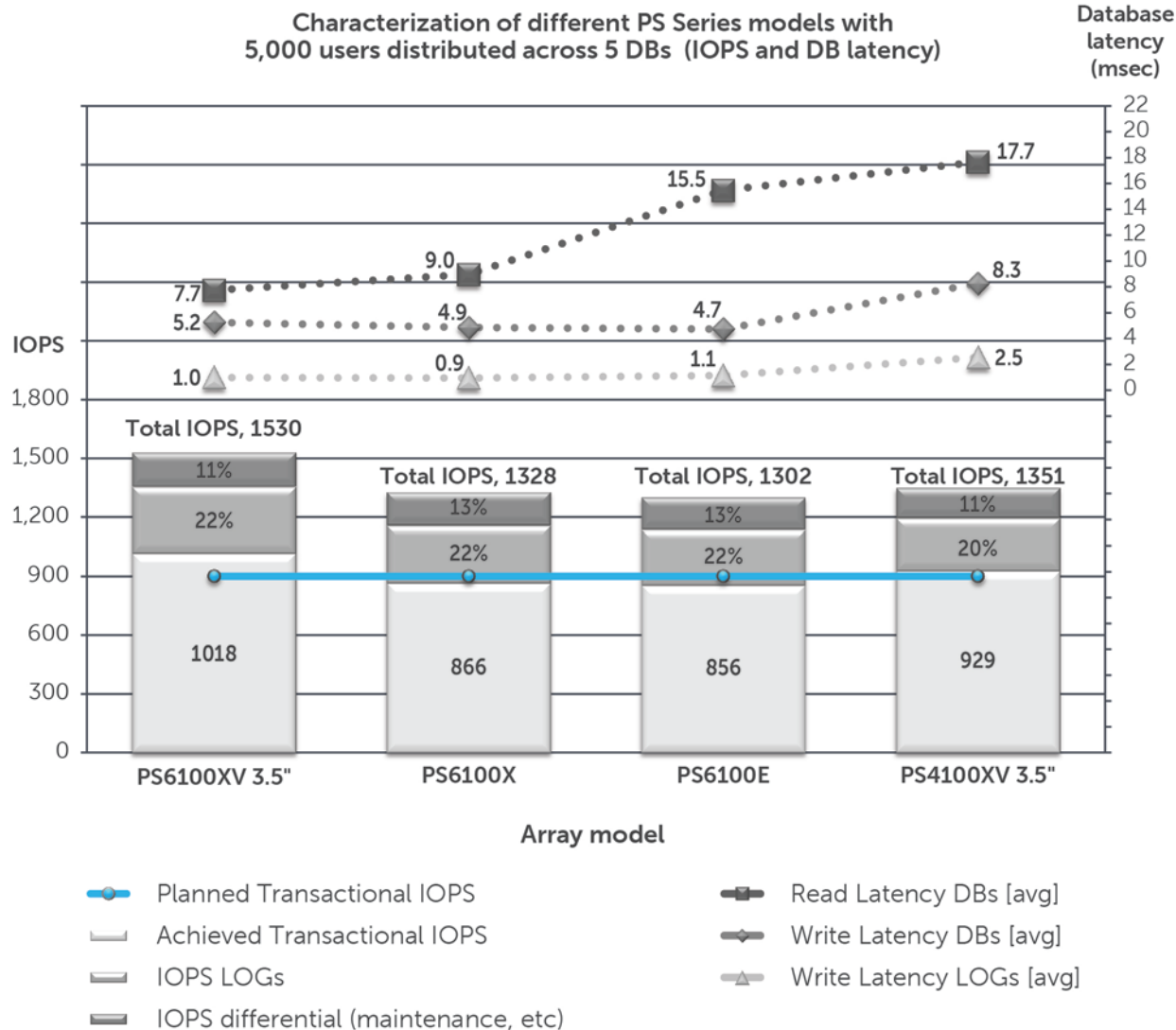


Figure 12 Characterization of different PS Series models with 5,000 users distributed across 5 DBs

Table 22 reports the relative improvement or decline of performance recorded while assessing the different EqualLogic array models from the PS6100 family and the PS4100XV 3.5" model. The percentages are calculated against the total IOPS shown in Figure 12, not only the transactional IOPS.

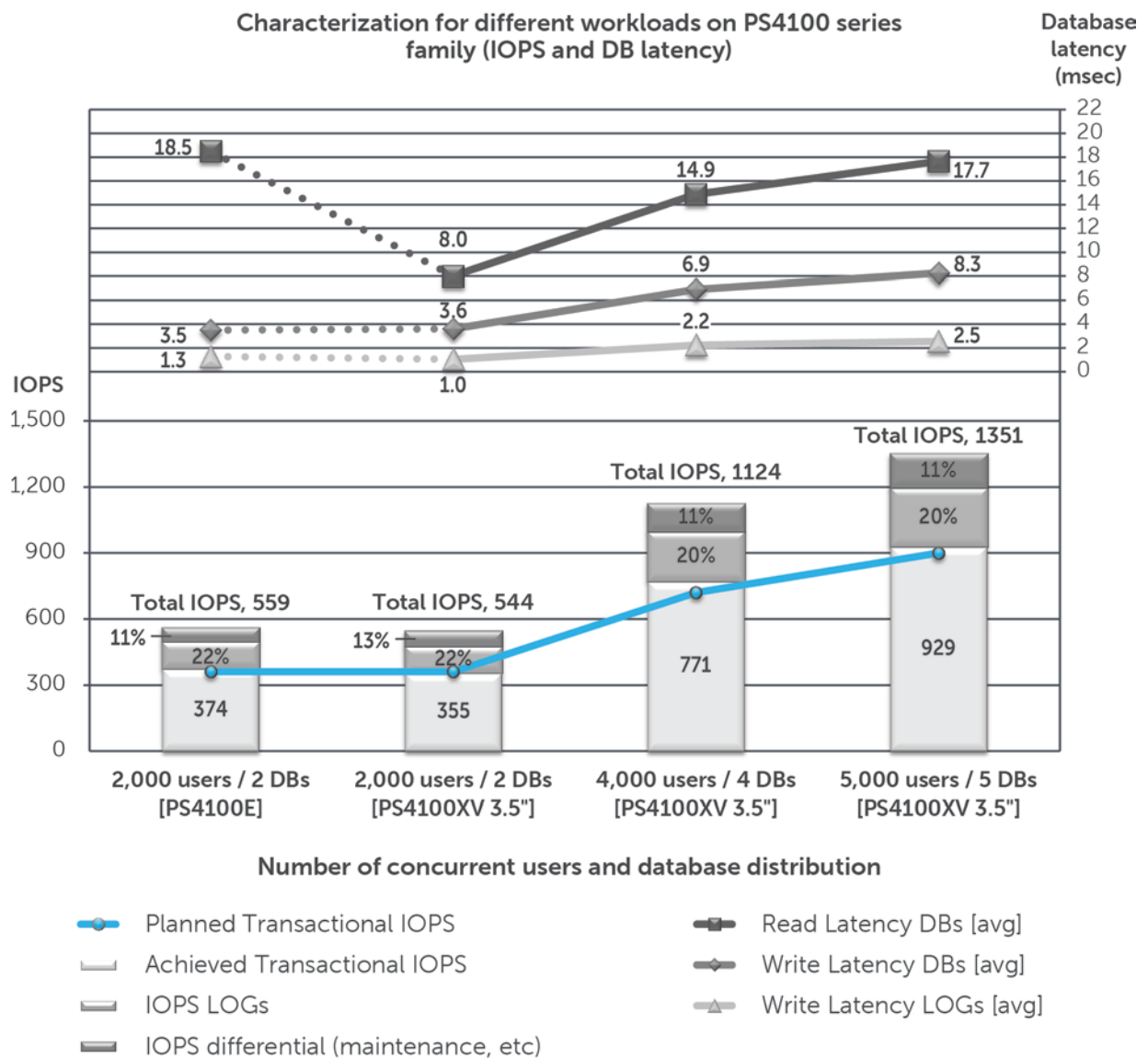
**Table 22 Test results: improvement or decline relationship of performance with different PS Series model**

|                      | <b>PS6100XV 3.5"</b> | <b>PS6100X</b>  | <b>PS6100E</b>   | <b>PS4100XV 3.5"</b> |
|----------------------|----------------------|-----------------|------------------|----------------------|
| IOPS planned         | 100% [900 IOps]      | 100% [900 IOps] | 100% [900 IOps]  | 100% [900 IOps]      |
| Total IOPS performed | 100% [1530 IOps]     | 87% [1328 IOps] | 85% [1302 IOps]  | 88% [1351 IOps]      |
| Read Latency DBs     | 100% [7.7 msec]      | 116% [9.0 msec] | 200% [15.5 msec] | 228% [17.7 msec]     |
| Write Latency DBs    | 100% [5.2 msec]      | 92% [4.9 msec]  | 90% [4.7 msec]   | 157% [8.3 msec]      |
| Write latency LOGs   | 100% [1.0 msec]      | 96% [0.9 msec]  | 118% [1.1 msec]  | 262% [2.5 msec]      |

The outcomes attest the higher rotational speed of the drives is the first aspect to factor for the performance. The PS6100XV 3.5" equipped with 15K RPM SAS drives leads the performance for the pool of models assessed, followed by the PS6100X with 10K RPM SAS drives, and then by the PS6100E with 7.2K NL-SAS drives, where the characteristic random IOPS of Exchange generate the ample gap within the read latency indicator. The PS4100XV 3.5" is capable of power the same 5,000 concurrent users of the parent family considering the 15K SAS drives, however it shows noticeable lower performance being the limited number of disk drives in the array, 12 versus the 24 of the PS6100 family (or 10 versus 22, seeing the arrays are configured in RAID 50 and the remaining disk drives are configured as Hot-Spares).

The results collected from the Exchange Jetstress simulation against the different PS4100 family models are reported in Figure 13.





**Figure 13** Characterization for different workloads on PS4100 series

Table 23 reports the relative improvement or decline of performance recorded while assessing the different EqualLogic array models from the PS4100 family. The percentages are calculated against the total IOPS shown in Figure 13, not only the transactional IOPS.

**Table 23** Test results: improvement or decline relationship of performance with different PS Series model

|                      | 2,000 users [PS4100E] | 2,000 users [PS4100XV 3.5"] | 4,000 users [PS4100XV 3.5"] | 5,000 users [PS4100XV 3.5"] |
|----------------------|-----------------------|-----------------------------|-----------------------------|-----------------------------|
| IOPS planned         | 39% [360 IOps]        | 39% [360 IOps]              | 80% [720 IOps]              | 100% [900 IOps]             |
| Total IOPS performed | 41% [559 IOps]        | 40% [544 IOps]              | 83% [1124 IOps]             | 100% [1351 IOps]            |
| Read Latency DBs     | 105% [18.5 msec]      | 45% [8.0 msec]              | 84% [14.9 msec]             | 100% [17.7 msec]            |
| Write Latency DBs    | 42% [3.5 msec]        | 44% [3.6 msec]              | 84% [6.9 msec]              | 100% [8.3 msec]             |
| Write latency LOGs   | 50% [1.3 msec]        | 40% [1.0 msec]              | 87% [2.2 msec]              | 100% [2.5 msec]             |

The outcomes of the characterization of the PS4100 family manifest a solid relationship between workload applied to the PS4100XV 3.5" and the resultant set of key latency indicators, resulting in a favorable predictability for different loads. The PS4100E, equipped with 7.2K NL-SAS drives, performs adequately, powering 2,000 concurrent users, but suffers like the equivalent parent model on read latency because of the random nature of the workload.

## 6 Capacity planning and sizing

The storage requirements to appropriately size Exchange Server 2010 mailbox role depend on two core factors: capacity and performance. The analysis presented in section 5 reveals the amount of variables surrounding and influencing these two factors.

### 6.1 Capacity considerations

Exchange capacity requirements are a summary of the elements previously mentioned in section 3: mailbox database files, transaction log files, checkpoint files and Search Catalog.

**Mailbox database capacity allocation** is calculated by adding the mailbox quota for each user, the database white space and the space allocated for single item recovery option, if enabled. The count of required capacity for an Exchange organization is summarized in the following series of formulas.

*Mailbox DB size = Number of mailboxes \* (Mailbox storage quota + Dumpster Delta + White Space)*

where

*Dumpster Delta = Dumpster size + Single Item Recovery + Calendar Logging*

*Dumpster size = #Sent & Received msgs per user per day \* Avg msg size \* Deleted item retention time*

*Single Item Recovery = (Mailbox storage quota \* 0.012)*

*Calendar Logging = (Mailbox storage quota \* 0.03)*

*White space = (Number of Sent & Received msgs per user per day \* Avg msg size)*

The assumptions for these formulas are the background database maintenance runs in background 24x7 (otherwise, without a full pass, the white space can exceed this estimate), and the 14 days default dumpster retention is maintained (or, recoverable items impact should be recalculated). Besides the dumpster delta allocation begins playing a role in using database space over the mailbox storage quota only when a mailbox is advancing its allocation near the mailbox quota threshold.

**Transaction log files capacity** is calculated estimating the number of transactions generated per day and the average size of messages. The log files in Exchange Server 2010 have a fixed size of 1 MB. Microsoft guidelines around log files generation weights each message for the size and adds an additional overhead originated from the descriptor tables of the message bodies and the attachments.

Additional factors to equate in the rate of log files generation are:

- move mailbox operations (across database or servers)
- data protection (backup) policies, including log files retention and recovery SLA
- high availability configuration, including log truncation options and space for lag log replay

Table 24 reassumes the numerical variables resulting from the Microsoft guidelines. A formula representation to provide the same outcome is reported here:

$$\text{Number of logs} = \frac{2.73 * \#Sent \ \& \ Received \ msgs \ per \ user \ per \ day * Avg \ msg \ size * msg \ size \ factor}{Size \ of \ log \ (1 \ MB)}$$

**Table 24 Microsoft guidelines for Exchange mailbox database log files generation**

| User usage profile | Number of logs per day per user | Average message size | Message size Factor | Descriptors overhead factor                                |
|--------------------|---------------------------------|----------------------|---------------------|--|
| 50 messages/day    | 10                              | 75 KB                | 1                   | $\frac{\text{Logs allocation}}{\#msg * msg \ size} = 2.73$ |
| 150 messages/day   | 30                              |                      |                     |  |
| 300 messages/day   | 60                              |                      |                     |  |
| 50 messages/day    | 19                              | 150 KB               | 1.9                 |  |
| 150 messages/day   | 57                              |                      |                     |  |
| 300 messages/day   | 114                             |                      |                     |  |
| 50 messages/day    | 38                              | 300 KB               | 3.8                 |  |
| 150 messages/day   | 114                             |                      |                     |  |
| 300 messages/day   | 228                             |                      |                     |  |

**Checkpoint files capacity** is ignored from the capacity assessment due to the irrelevant size (8 KB).

**Search Catalog capacity** (or Content Index) is usually calculated as 10% of the mailbox database size.

**Note:** For additional information about the Exchange 2010 capacity requirements refer to Microsoft documentation *Understanding Mailbox Database and Log Capacity Factors*, available at: <http://technet.microsoft.com/en-us/library/ee832796.aspx>

**EqualLogic usable capacity** represents the amount of space available in the SAN after the selection of array family and model, disk drives size and RAID policy (which also defines the number of spare drives as described in section 5.5). Table 25 displays the different drives alternatives available when selecting the arrays. The table is not exhaustive as it does not include every EqualLogic array model and the available drives size is subject to change overtime.

**Table 25 EqualLogic PS series drives and usable capacity examples (1 TB = 10<sup>3</sup> GB)**

| PS series family           | Disk drive size & array model | Capacity raw | Capacity RAID 10 | Capacity RAID 50 | Capacity RAID 5 | Capacity RAID 6 |
|----------------------------|-------------------------------|--------------|------------------|------------------|-----------------|-----------------|
| PS6100<br>(24 disk drives) | 600 GB, XV 3.5"               | 14.4 TB      | 6600 GB          | 10800 GB         | 12600 GB        | 11400 GB        |
|                            | 300 GB, X 2.5"                | 7.2 TB       | 3300 GB          | 5400 GB          | 6300 GB         | 5700 GB         |
|                            | 600 GB, X 2.5"                | 14.4 TB      | 6600 GB          | 10800 GB         | 12600 GB        | 11400 GB        |
|                            | 900 GB, X 2.5"                | 21.6 TB      | 9900 GB          | 16200 GB         | 18900 GB        | 17100 GB        |
|                            | 1 TB, E 3.5"                  | 24 TB        | 11000 GB         | 18000 GB         | 21000 GB        | 19000 GB        |
|                            | 2 TB, E 3.5"                  | 48 TB        | 22000 GB         | 36000 GB         | 42000 GB        | 38000 GB        |
|                            | 3 TB, E 3.5"                  | 72 TB        | 33000 GB         | 54000 GB         | 63000 GB        | 57000 GB        |
| PS4100<br>(12 disk drives) | 600 GB, XV 3.5"               | 7.2 TB       | 3000 GB          | 4800 GB          | 6000 GB         | 5400 GB         |
|                            | 1 TB, E 3.5"                  | 12 TB        | 5000 GB          | 8000 GB          | 10000 GB        | 9000 GB         |
|                            | 2 TB, E 3.5"                  | 24 TB        | 10000 GB         | 16000 GB         | 20000 GB        | 18000 GB        |
|                            | 3 TB, E 3.5"                  | 36 TB        | 15000 GB         | 24000 GB         | 30000 GB        | 27000 GB        |

## 6.2 Performance considerations

Exchange performance requirements are a characterization of the kind of IOPS generated by the Exchange mailbox role components and the average user profile assessed for a particular organization.

**Mailbox database performance** is featured by the responsiveness of the storage subsystem in serving the IO requests. The database transactional access is purely random, with a combination of reads and writes depending on the user profile as shown in section 5.1. The database block size of 32 KB drives a medium to small size of IO. The checksum activity is sequential in nature and performed with 64 KB reads. The database file is subject to large sequential IO (reads) in case of backup operations or database replica initialization (DAG seeding process).

**Transaction log files performance** is derived as the mailbox database from the storage subsystem pace of reaction when accessed. The log files are written sequentially as soon as a transaction is closed with a small size of IO, after have considered the log checkpoint depth behavior exposed in section 3.1. The log files are subject to a sequential read activity in presence of DAG configuration or during a replay process in case of database recovery.

**Checkpoint files performance** is ignored from the performance assessment due to the limited IOPS expected and their size (8 KB).

**Search Catalog** (or Content Index) is accessed randomly with both read and write methods. The content index is a collection of flat files of different sizes with an expected medium size of IO.

**EqualLogic performance planning** characteristic is based on a list of factors discussed in section 5.5 and summarized here:

- **Number of disk drives** in the array is defined by the PS Series family selected: 24 drives for the PS6100 and 12 drives for the PS4100. More spindles available offer an improved support for applications with a storage access multi-threaded as Exchange.

- **Disk drives rotational speed:** the more RPM are supplied by the drive platters the faster the IOPS requested from the host are served. The disk drive speeds currently are 15K SAS, 10K SAS and 7.2K NL-SAS (SSD drives are omitted from the list as they offer a price/performance ratio and level of IOPS inappropriate for Exchange 2010 requirements).
- **RAID policy** offers data protection from drives failure. Each RAID level provides a characteristic behavior depending on the workload applied (sequential versus random, small versus large IO, reads versus writes). The current RAID policies implemented in the EqualLogic array families assessed are RAID 10, 50, 5 and 6, where RAID 10 provides the best performance due to the absence of parity calculation when compared with the three remaining policies.

### 6.3 Sizing example

The example provided assumes we are engaged to assess and design a storage solution for the messaging infrastructure of an imaginary organization with the following requirements:

- A company with strong growth expectation
- 6,000 skilled messaging users connecting mainly from headquarters or the home office to the datacenter at headquarters, with an assessed average usage profile of 200 messages per day and a mailbox storage quota requirement of 700 MB
- 24/7 rotational schedule to support a global call center
- A highly available access with a nearly instantaneous recovery in case of data/server failure
- Needs to integrate with a pre-existing disaster recovery plan including a secondary datacenter

The first level of design expects to deploy the following components is:

- An estimated company growth factor of 15%, equating to a total user of 6,900
- Four Exchange mailbox servers configured within one DAG
- One mailbox server with all the active databases, all the remaining with replica copies
- Two servers per each datacenter
- 1,000 mailboxes per mailbox database, a total of 7 active databases
- Background maintenance enabled 24/7
- A growth overhead factor of at least 20%
- An estimated transactional IOPS requirement of 1656 IOPS (0.24 per mailbox)
- We assume an average message size of 75 KB, a default deleted item retention of 14 days and decided to have single item recovery enabled (because of the trained users)

**Exchange Server cache (estimated)** minimum requirement is calculated by

$$\text{Memory} = \text{min RAM with 1 to 10 DBs} + (\#users * \text{Cache per user}) = 2 \text{ GB} + (6900 * 12 \text{ MB}) = 85 \text{ GB}$$

and requires a building block server or virtual machine with 96 GB of RAM.

**Mailbox databases capacity size** is calculated by

$$\text{Dumpster size} = (200 \text{ messages per day} * 75 \text{ KB} * 14 \text{ days}) = 210 \text{ MB}$$

$$\text{Single Item Recovery} = (700 \text{ MB} * 0.012) = 8.4 \text{ MB}$$

$$\text{Calendar Logging} = (700 \text{ MB} * 0.03) = 21 \text{ MB}$$

$$\text{Dumpster Delta} = 210 \text{ MB} + 8.4 \text{ MB} + 21 \text{ MB} = 240 \text{ MB}$$

*White space* = (200 messages per day \* 75 KB) = 15 MB

*Mailbox DB size* = 6900 users \* (700 MB + 240 MB + 15 MB) = 6900 \* 955 MB = 6590 GB

**Transaction log files capacity** is calculated by

$$\text{Number of daily logs} = 6900 * \frac{2.73 * 200 \text{ messages per day} * 75 \text{ KB} * 1}{1 \text{ MB}} = 6900 * 41 = 282900$$

We are not aware of the data protection solution implemented. We plan for three days log retention on the volume assuming log files backup and truncation happens on a daily basis.

*Estimated Log files capacity* = Number of daily logs \* Log Retention time = 282900 \* 3 = 849 GB

**Search Catalog capacity size** is calculated by

*Search Catalog* = (0.10 \* Mailbox DB size) = 659 MB

**Total capacity** required is calculated by

*Total capacity* = (Mailbox Databases + Log files + Search Catalog) = (6590 + 849 + 659) = 8098GB

**Volume capacity** required is calculated by

$$\text{Volume capacity} = \frac{\text{Total capacity}}{\text{Number of databases}} * (1 + \text{overhead growth factor}) = \frac{8098}{7} * 1.20 = 1388GB$$

**Performance planning** required is calculated by

*Transactional IOPS* = #of users \* avg IOPS per user \* DAG reduction factor = 6900 \* 0.20 = 1380 IOPS

Imputing the results from section 5.1 in the scenario of 0.18 IOPS per mailbox the IOPS required for the log files access are around 22% and the differential for maintenance 11% of the total IOPS with the figures decreasing while increasing the usage profiles. We estimate for 0.20 IOPS per mailbox the percentage are very close at 20% and 10%, leaving 70% for the transactional IOPS.

$$\text{Total IOPS (equated from Transactional)} = \frac{\text{Transactional IOPS}}{0.7} = \frac{1380}{0.7} = 1972 \text{ IOPS}$$

The size of the database around 1 TB, according to the outcomes of section 5.2, does not affect IOPS, but slightly the read latency which we should account for.

The database layout of 7 databases requires an adjustment of 1-2% more for the log files IOPS, and around 3% for the IOPS differential (maintenance, etc), according to the finding of section 5.3.

The DAG configuration with 4 database replica copies would require an additional 3% of log files IOPS due to the additional reads, according to section 5.4.

The resultant set of IOPS is:

*Log IOPS* = Total IOPS \* (0.2 + DAG diff + DB layout) = 1972 \* (0.2 + 0.03 + 0.02) = 493 IOPS

*Diff IOPS (maintenance, etc)* = Total IOPS \* (0.1 + DB layout) = 1972 \* (0.1 + 0.03) = 257 IOPS

*Total IOPS (final)* = Transactional + Logs + Diff(maintenance, etc) = 1972 + 493 + 257 = 2722 IOPS

**EqualLogic SAN selection** requires the evaluation of capacity and performance. Our capacity requirement is 9800 GB of total usable space for each of the four servers, and 2722 IOPS per server, while we deploy in two different physical locations.

The RAID policy proposed is RAID 6 due to the requirements of high resiliency of the solution and it requires a minimal increase in read latency when compared to the other policies.

According to the finding of section 5.8, in order to run 6,900 concurrent users with an usage profile slightly higher than what we assessed, the selection available is restricted to two models (PS6100XV 3.5" or PS6100X) if we plan one unit per server, or four models in case we decide to dedicate two units per server instead (also PS6100E or PS4100XV 3.5").

The total IOPS required are approximately in the range of the test case in section 5.6 with the PS6100XV 3.5". The capacity requirements matched with Table 25 in section 6.1 confirm the array, disk drives and RAID level provides room for the solution.

**Table 26 Sizing example summary**

| <b>Requirements</b>                             |   |
|---|---|
| Number of mailboxes/users                       | 6,000 concurrent users  |
| Company growth expectation                      | 15%   |
| Messages sent/received per day per mailbox      | 200 messages  |
| High availability                               | Highly resilient  |
| Usage pattern                                   | 24/7  |
| Deployment location                             | Multi sites   |
| Assumptions                                     | Average message size 75 KB<br>Deleted item retention 14 days  |
| <b>Proposed solution</b>                        |   |
| Number of mailboxes/users                       | 6,900 concurrent users  |
| Mailbox size                                    | 700 MB  |
| Number of databases                             | 7 databases (active)  |
| Number of database replica copies               | 4 nodes DAG, 21 database replica copies   |
| Databases size                                  | 942 GB each   |
| Log files size                                  | 41 GB per database with 3 days retention  |
| Search Catalog                                  | 95 GB per database  |
| Volumes   | 7 volumes (DB + Logs+ Search Catalog)<br>1400 GB each (>120% of the planned data)                                 |
| Mailbox allocation                              | 857 to 986 mailboxes per each mailbox database (6,000 to 6,900 users)   |
| IOPS per mailbox / messages per day per mailbox | 0.24 IOPS / 200 messages<br>reduced to 0.20 IOPS per DAG factor   |
| RAID policy                                     | RAID 6 for best availability  |
| Array model, amount of units, SAN configuration | 4x PS6100XV 3.5", 24x 600 GB disk drives<br>2 groups, one per each site<br>4 pools, 2 per site, 1 per each server |
| Background database maintenance                 | Enabled   |

## 7 Best practices recommendations

Refer to these best practices to plan and configure Exchange Server 2010 and EqualLogic arrays.

### Storage best practices

- Use MPIO DSM provided by EqualLogic HIT Kit to improve performance and reliability for the iSCSI connections.
- Maintain a 1:1 ratio between the number of network ports on the active arrays controller and the number of host network adapters to utilize the available bandwidth.
- Choose carefully the most appropriate RAID policy from the beginning according with the performance, capacity, and tolerance to failure requirements of your environment.
- RAID conversion provides outstanding benefits when strictly required. Do not overuse it if a clean configuration is viable. RAID conversions of large volumes take a considerable amount of time and leave the resultant set of data pages extremely fragmented with a potential threat of lower performance.
- Do not share the disk drives for active and replicated copies of an Exchange mailbox database. The failure of a set of drives with multiple database copies decreases the resilience or the perceived availability in this deployment scenario. Dedicate separate pools for databases connected to different nodes in a DAG instead.
- Distribute network port connections on the controllers accordingly with the port failover mechanism and the redundancy implemented on the network switches.

### Network best practices

- Design separate network infrastructures to isolate the LAN traffic from the SAN traffic (iSCSI)
- Implement redundant components (switches, ISLs, network adapters) to provision a resilient network infrastructure between the endpoints (stack, LAG, load balancing or network card teaming)
- Enable flow control for the switch ports hosting the PS Series array controller connections
- Enable flow control for the network adapters dedicated to SAN traffic (iSCSI)
- Disable spanning tree from the switch ports hosting the PS Series array controllers connections, and enable 'Portfast' instead
- Enable jumbo frames (large MTU) for the switch ports hosting the PS Series array controller connections
- Evaluate jumbo frames (large MTU) for the LAN network when appropriate (limited by the type of devices the traffic traverses)

### ESX Hypervisor and virtual machine best practices

- Design separated virtual switches to isolate the Virtual Machines LAN traffic from the SAN traffic (iSCSI).
- Configure Port Group and VLAN tagging to logically segregate different LAN traffic.
- Enable jumbo frames (large MTU) for the virtual switches assigned to SAN traffic (iSCSI).
- Evaluate jumbo frames (large MTU) for the virtual switches assigned to LAN traffic.
- Reserve at least two network adapters for each virtual switch to achieve resiliency.
- Maintain a 1:1 ratio between the number of network ports on the active arrays controller and the number of virtual network adapter configured in the virtual machines.



- Configure a dedicated Port Group for each virtual network adapter connected to the SAN traffic you plan to have in the virtual machines, and add one physical network adapter for every Port Group isolating it by the override switch failover option. To provide failover capability, add other network adapters present in the virtual switch to the standby list in the Port Group.
- Select performance optimized network adapters of VMXNET3 type for guest network adapters connected to the SAN network (VMware tools required in the guest operating system).
- Enable TSO (TCP Segmentation Offload) and LRO (Large Receive Offload) on the guest network adapters connected to the SAN network (iSCSI).

### **Exchange installation best practices**

- Use Basic disk type for all EqualLogic volumes.
- Use GUID partition table (GPT) for Exchange volumes.
- Use default disk alignment provided by Windows 2008 or greater.
- Use NTFS file system with 64 KB allocation unit for Exchange database and log partitions.
- Deploy Windows operating system and Exchange data in physically separated disk drives.
- Database and log file isolation is not required when deployed in a DAG environment.
- Leave background database maintenance enabled (24x7) and account for the additional load. The BDM is anyway activated by default on every replica copy of your DAG configuration.

### **Know your workload**

A sizing plan is viable only when enough information of the current user activities is provided. In contrast, for a greenfield deployment, estimates based on the business case can be used to define the average usage profiles.

### **Exchange Capacity sizing**

Do not simplify the calculations required to understand the capacity requirements of a solution. Use the guidelines provided in this paper and the Microsoft guidelines to properly assess the storage capacity requirement for database and log files, thus of the SAN.

### **Growth and overhead**

Understand your wider business requirements and company strategy to prevent organization growth. Design your environment based on building blocks easily replicable to face sudden changes in business requirements.

Always account for at least 20% overhead of the maximum sizes calculated to address unforeseen growth.

### **Exchange mailbox database layout**

The balance between having few or more databases for the same amount of mailbox users depends on administrative considerations. Bigger databases fitted with a great amount of users perform better than an extensive number of small databases because of the aggregation of overhead loads, but a widespread number of databases become counterproductive for administrative purposes.

Microsoft Windows operating system supports a precise maximum amount of iSCSI targets and connections: do not overcome the limits to avoid an unsupported deployment scenario.

## Number of mailboxes per mailbox database

Reducing the number of users per database provides a more agile size to administer in case of using a traditional backup application or even when administrative tasks require to temporarily dismount the database causing a downtime. The balance between number of databases and users per databases must be cautiously planned according with the needs and size of the entire Exchange organization. Usually a single database containing a high percentage of the entire organization is perceived as a single point of failure or at least as an administrative constraint.

Note: For general recommendations and information about EqualLogic PS Series array configurations, refer to Dell EqualLogic Configuration Guide, available at:  
<http://www.delltechcenter.com/page/EqualLogic+Configuration+Guide>

# Appendix A Configuration details

## A.1 Hardware components

Table 27 lists the details of the hardware components used for the configuration setup.

**Table 27 Hardware components**

| <b>Test configuration – Hardware components</b> |  |
|---|--|
| <b>Servers</b>                                  | <ul style="list-style-type: none"> <li>• Dell PowerEdge M1000e Blade enclosure               <ul style="list-style-type: none"> <li>○ 2x Chassis Management Controller (CMC), Firmware 3.21.A00</li> </ul> </li> <li>• Dell PowerEdge M710 Blade Server               <ul style="list-style-type: none"> <li>○ 2x Quad Core Intel Xeon X5570 Processors, 2.93 Ghz, 8M Cache</li> <li>○ RAM 64 GB (8x 8GB)</li> <li>○ PERC 6/I RAID controller</li> <li>○ 2x 73 GB 15K SAS (RAID-1) + 2x 146 GB 15K SAS (RAID-1)</li> <li>○ 4x Broadcom NetXtreme II 5709S Dual Port 1GbE Mezzanine Card, Firmware 6.4.5</li> </ul> </li> <li>• Dell PowerEdge M610 Blade Server               <ul style="list-style-type: none"> <li>○ 2x Quad Core Intel Xeon E5520 Processors, 2.26 Ghz, 8M Cache</li> <li>○ RAM 32 GB (4x 8GB)</li> <li>○ PERC 6/I RAID controller</li> <li>○ 2x 73 GB 15K SAS (RAID-1)</li> <li>○ 2x Broadcom NetXtreme II 5709S Dual Port 1GbE Mezzanine Card, Firmware 6.4.5</li> </ul> </li> </ul>  |
| <b>Network</b>                                  | <ul style="list-style-type: none"> <li>• 2x Dell PowerConnect M6220 Blade Switches (stacked), Firmware 4.1.1.9               <ul style="list-style-type: none"> <li>○ installed in M1000e blade enclosure fabrics A1 and A2</li> </ul> </li> <li>• 2x Dell PowerConnect M6348 Blade Switches (stacked) , Firmware 4.1.1.9               <ul style="list-style-type: none"> <li>○ installed in M1000e blade enclosure fabrics B1 and B2</li> </ul> </li> <li>• 2x Dell PowerConnect 7048R Switches (stacked) , Firmware 4.1.1.9               <ul style="list-style-type: none"> <li>○ Installed top of the rack</li> </ul> </li> </ul>   |
| <b>Storage</b>                                  | <ul style="list-style-type: none"> <li>• 3x Dell EqualLogic PS6100XV 3.5"               <ul style="list-style-type: none"> <li>○ Dual 4-port 1GbE controllers, Firmware 5.1.2</li> <li>○ 24x 600GB 15K 3.5" SAS disk drives, raw capacity 14.4 TB each</li> </ul> </li> <li>• 1x Dell EqualLogic PS6100X               <ul style="list-style-type: none"> <li>○ Dual 4-port 1GbE controllers, Firmware 5.1.2</li> <li>○ 24x 900GB 10K 2.5" SAS disk drives, raw capacity 21.6 TB</li> </ul> </li> <li>• 1x Dell EqualLogic PS6100E               <ul style="list-style-type: none"> <li>○ Dual 4-port 1GbE controllers, Firmware 5.1.2</li> <li>○ 24x 2TB 7.2K 3.5" NL-SAS disk drives, raw capacity 48 TB</li> </ul> </li> <li>• 1x Dell EqualLogic PS4100XV 3.5"               <ul style="list-style-type: none"> <li>○ Dual 2-port 1GbE controllers, Firmware 5.1.2</li> <li>○ 12x 600GB 15K SAS disk drives, raw capacity 7.2 TB</li> </ul> </li> <li>• 1x Dell EqualLogic PS4100E               <ul style="list-style-type: none"> <li>○ Dual 2-port 1GbE controllers, Firmware 5.1.2</li> <li>○ 12x 2TB 7.2K NL-SAS disk drives, raw capacity 24 TB</li> </ul> </li> </ul> |

## A.2 Software components

The setup of the environment required the deployment of the following software components:

- Bare-metal hypervisor: VMware 5.0 ESXi on all host servers
- Operating System: Windows Server 2008 R2 on all virtual machines
- VMware vCenter Server, Client and CLI for the virtual infrastructure management (vCENTER virtual machine)
- EqualLogic Host Integration Toolkit to access the storage SAN with MPIO
- EqualLogic SAN Headquarters to monitor the health and performance of SAN
- Microsoft Exchange Jetstress to simulate the access to the storage subsystem (all MBX virtual machines)

The following elements were installed and configured to simplify the management of the environment, but not required for the goal of the tests:

- Active Directory Domain Services and DNS Server roles for the domain controller (DC01 virtual machine)

Table 28 lists the details of the software components and version used for the configuration setup.

**Table 28 Software components**

| <b>Test configuration – Software components</b> |  |
|---|--|
| <b>Operating Systems</b>                        | <ul style="list-style-type: none"> <li>• Host: VMware ESXi 5.0.0 (build 469512)</li> <li>• Guests: Microsoft Windows Server 2008 R2 Enterprise Edition Service Pack 1 (build 7601)               <ul style="list-style-type: none"> <li>○ Dell EqualLogic Host Integration Toolkit 3.5.1</li> <li>○ MPIO enabled using EqualLogic DSM for Windows</li> <li>○ VMware tools 8.6.0</li> </ul> </li> </ul> |
| <b>Applications</b>                             | <ul style="list-style-type: none"> <li>• VMware vSphere 5.0.0 vCenter Server and Client (build 456005)</li> <li>• VMware vSphere 5.0.0 CLI (build 435426)</li> </ul>   |
| <b>Monitoring tools</b>                         | <ul style="list-style-type: none"> <li>• Dell EqualLogic SAN Headquarters 2.2 (build 2.2.0.5924)</li> <li>• Microsoft Performance Monitor (built in the Windows Operating System)</li> </ul>   |
| <b>Simulation tools</b>                         | <ul style="list-style-type: none"> <li>• Microsoft Exchange Jetstress 2010 (build 14.01.0225.017)</li> <li>• Microsoft Exchange 2010 Server Database Storage Engine and Library Service Pack 2 (build 14.01.0218.012)</li> </ul>   |

### A.3 Network configuration details

Two physical networks were built in order to provide full isolation between regular IP traffic and iSCSI data storage traffic. Also, each IP network was segregated from the others by the use of VLANs with tagged traffic. In order to achieve network resiliency for hardware faults, we stacked at least two physical switches for each network, as well as made redundant the uplink (LAG) between the stacked switches. Some configurations worth mentioning were:

- Flow control enabled for every switch port on 7048Rs and M6348s
- Spanning tree Portfast enabled for every switch port on 7048Rs and M6348s
- Jumbo frames enabled for every switch port on 7048Rs and M6348s

Table 29, Table 30, and Table 31 summarize the different networks required for the configuration setup and their usage.

**Table 29 Configuration – Switch modules**

| Switch Module         | Placement            | Purpose                    |
|-----------------------|----------------------|----------------------------|
| PowerConnect M6220 #1 | M1000e I/O Module A1 | Regular IP traffic         |
| PowerConnect M6220 #2 | M1000e I/O Module A2 | Regular IP traffic         |
| PowerConnect M6348 #1 | M1000e I/O Module B1 | iSCSI data storage traffic |
| PowerConnect M6348 #2 | M1000e I/O Module B2 | iSCSI data storage traffic |
| PowerConnect 7048R #1 | Top of the rack      | iSCSI data storage traffic |
| PowerConnect 7048R #2 | Top of the rack      | iSCSI data storage traffic |

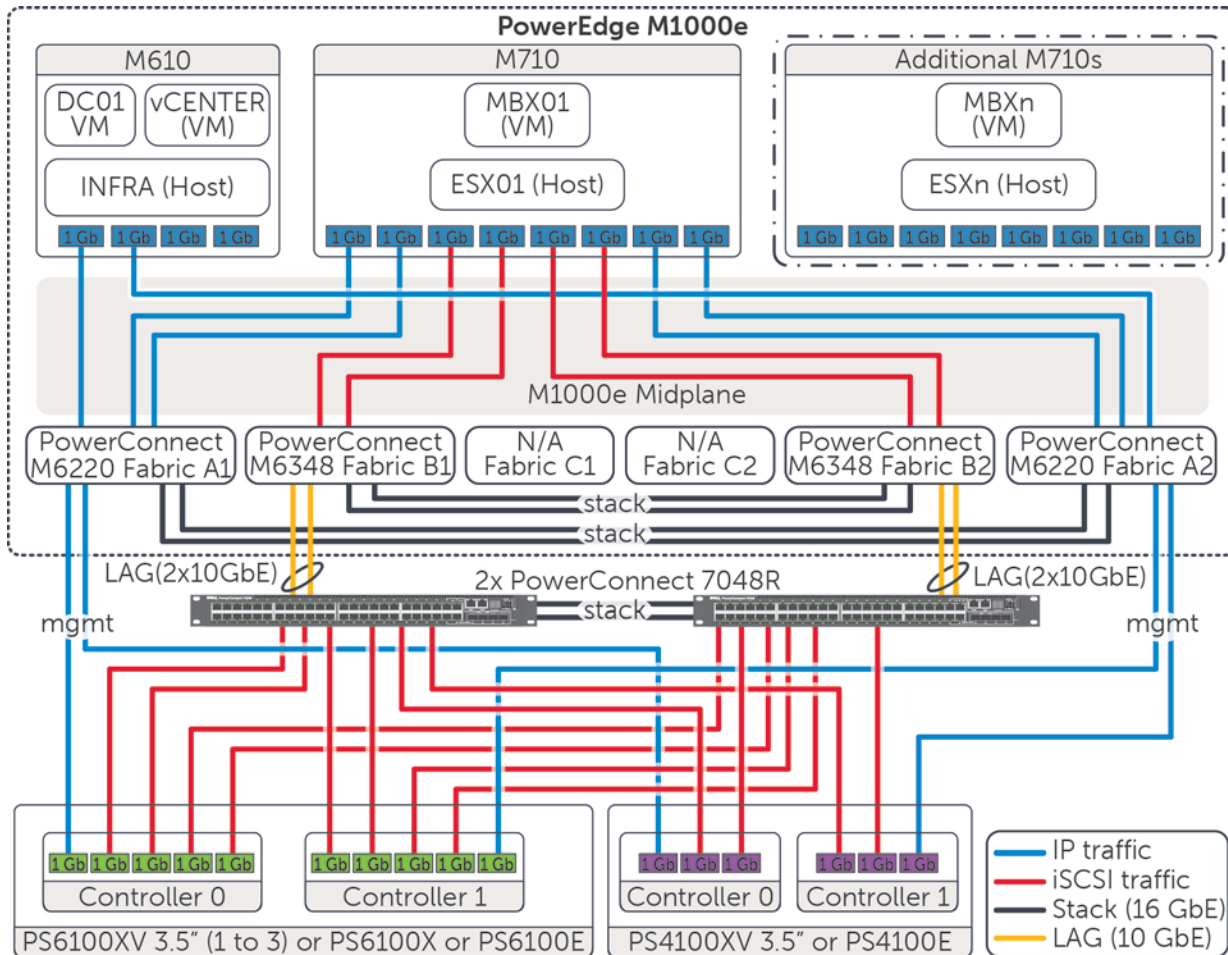
**Table 30 Configuration – Hosts to switch connection**

| Server         | Interface                                      | Number of NIC ports | Purpose                    |
|----------------|--|---------------------|----------------------------|
| PowerEdge M710 | LOM on A1                                      | 2                   | Regular IP traffic         |
|                | LOM on A2                                      | 2                   |                            |
|                | Mezzanine Card on B1                           | 2                   | iSCSI data storage traffic |
|                | Mezzanine Card on B2                           | 2                   |                            |
|                | Total ports = 8 (4 on Fabric A, 4 on Fabric B) |                     |                            |
| PowerEdge M610 | LOM on A1                                      | 1                   | Regular IP traffic         |
|                | LOM on A2                                      | 1                   |                            |
|                | Mezzanine Card on B1                           | 1                   | iSCSI data storage traffic |
|                | Mezzanine Card on B2                           | 1                   |                            |
|                | Total ports = 4 (2 on Fabric A, 2 on Fabric B) |                     |                            |

**Table 31 Configuration – VLANs**

| VLAN ID     | Switch it is implemented on            | Purpose                      |
|-------------|--|------------------------------|
| 100         | PowerConnect M6220                     | Management (Service Console) |
| 200         | PowerConnect M6220                     | LAN traffic                  |
| 1 (default) | PowerConnect M6348, PowerConnect 7048R | iSCSI                        |

Figure 14 presents the diagram of network connections between the blade servers and the storage arrays.



**Figure 14 Network connectivity diagram**

## A.4 Host hypervisor and virtual machines configuration

A virtual infrastructure built with VMware vSphere hosted all the components of the test infrastructure. Some key elements of the virtual infrastructure configuration were:

- VMware 5.0 ESXi deployed on all hosts, and centrally managed by the same vCenter server
- Up to 3 hypervisor hosts (M710s) configured as standalone server
- 1 hypervisor host (M610) configured as standalone
- All guests deployed from a single template
- Guest VM OS disks statically deployed into the local VMFS data store of its respective host
- Guest iSCSI initiator used to access volumes hosted on the EqualLogic SAN

Table 19 lists the relation between the hypervisor host and each virtual machine, with a brief summary of the virtual resources allocated for each virtual machine.

**Table 32 Configuration – Host and guest allocation**

| Host         | VM      | Purpose  | vCPUs | Memory | Network Adapters |
|--------------|---------|--|-------|--------|------------------|
| INFRA (M610) | DC01    | Active Directory Domain Controller                 | 2     | 4GB    | 1x E1000         |
|              | vCENTER | VMware vCenter                                     | 2     | 4GB    | 2x E1000         |
| ESXn (M710s) | MBXn    | Exchange Server Maibox role simulation (Jetstress) | 2     | 16GB   | 1x E1000         |
|              |         |  |       |        | 4x VMXNET3       |

### A.4.1 Virtual network configuration

The virtual networks configured on the ESXi hypervisor servers were configured following the guidelines listed below.

- One virtual switch aggregating all the network adapters designated for regular IP traffic, with the default load balancing policy ('Route based on the originating virtual switch port ID')
  - 'Management Network' port group (VMkernel type) – all hosts
  - 'vCenters' port group (Virtual Machine type) – only for the INFRA host
  - 'LAN Traffic' port group (Virtual Machine type) – all hosts
  - Each port group traffic segregated by tagged packets (VLANs)
- One virtual switch aggregating all the network adapters designated for iSCSI storage traffic, with the default load balancing policy ('Route based on the originating virtual switch port ID')
  - One 'iSCSI Traffic' port group (Virtual machine type) for each physical network adapter assigned to the virtual switch – only for SAN attached virtual machines
  - Override the switch failover default configuration for each port group, allowing only one active adapter in each port group (on a rotational basis) and adding the remaining adapters in the standby list to maintain failover ability – only for SAN attached virtual machines
  - The network adapter failover order addresses the recover from a mezzanine card adapter failure first and then from a switch fabric failure.
- Jumbo frames enabled on both virtual switches
- Flow control on received and transmitted traffic (RX/TX) enabled on each physical network adapter

The vSwitch0 and vSwitch1 layout on the ESX01 host are shown in Figure 15 and Figure 16.

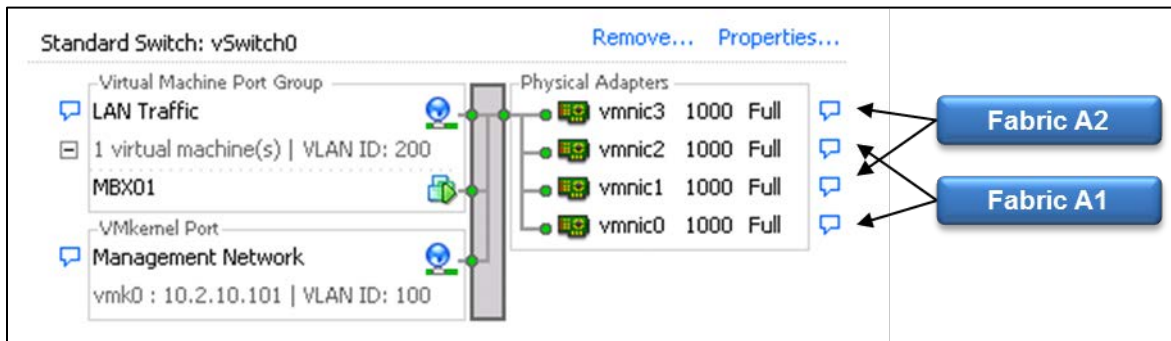


Figure 15 Configuration – vSwitch0 on ESX01

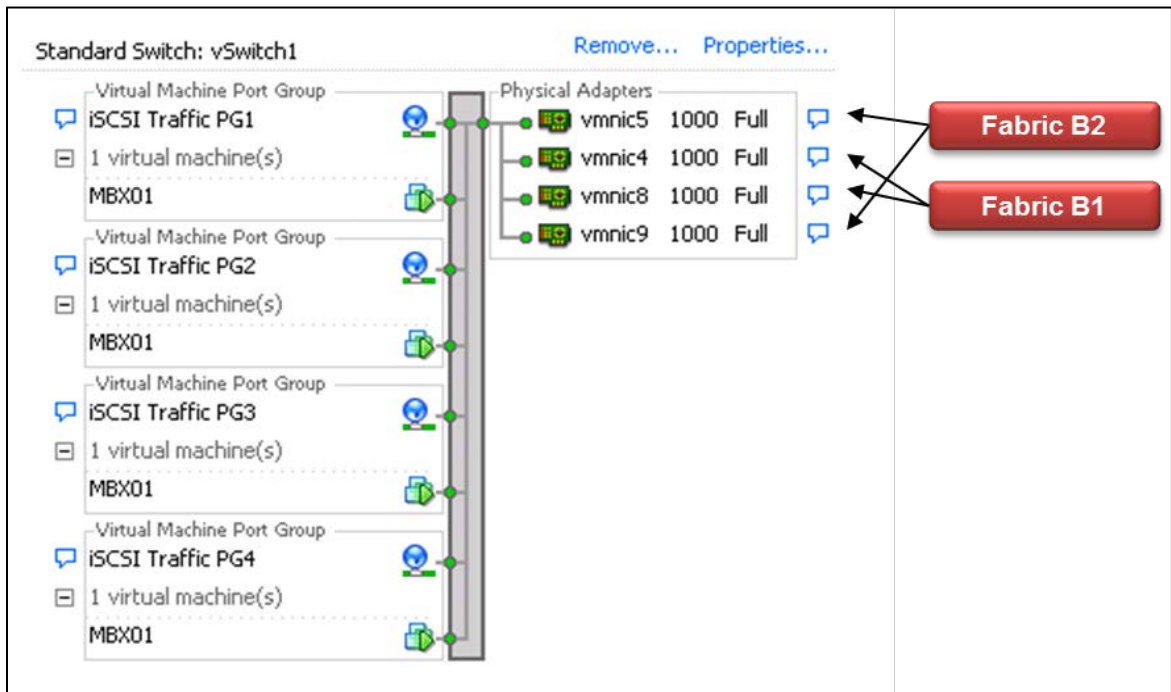


Figure 16 Configuration – vSwitch1 on ESX01

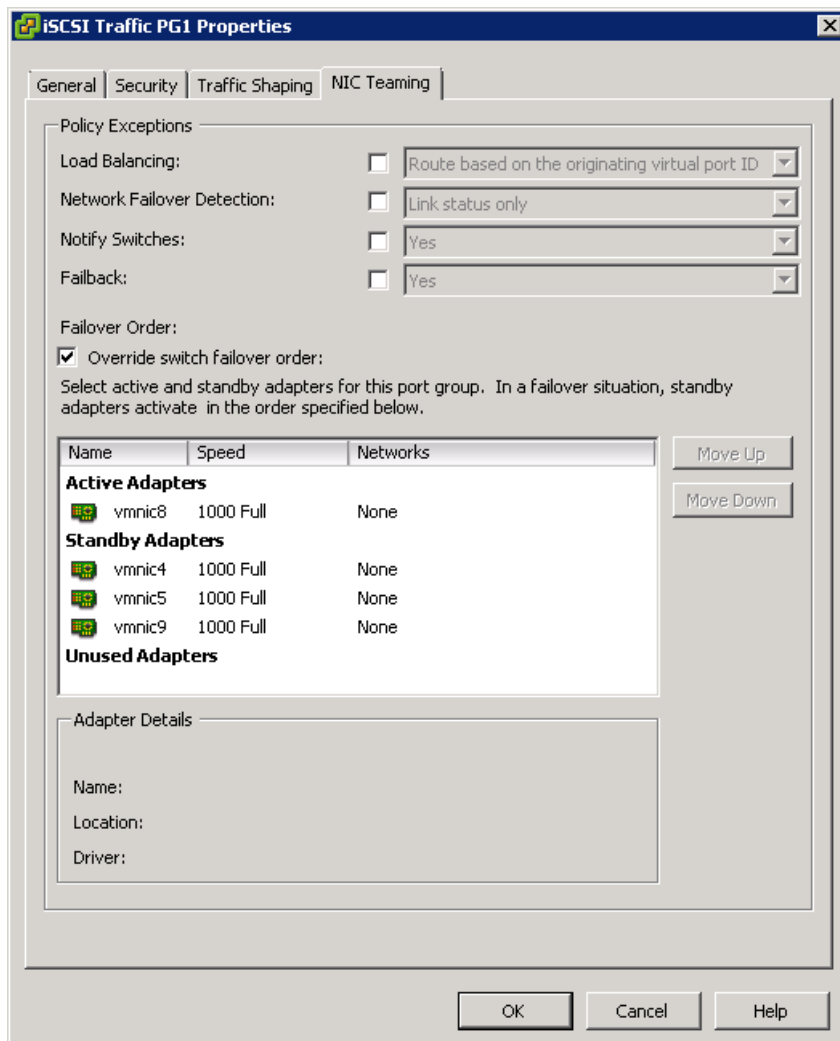
Table 33 reports the relationship between virtual switches, network adapters, and VLANs for each hypervisor host.



**Table 33 Configuration – Virtual switches, port groups, and network adapters**

| Host                            | Virtual Switch   | Network Adapters                | Port Group       | VLAN ID |
|---------------------------------|------------------|---------------------------------|------------------|---------|
| INFRA (M610)                    | vSwitch0         | vmnic0, vmnic1                  | Management       | 100     |
|                                 |                  |                                 | vCenter          | 100     |
|                                 |                  |                                 | LAN Traffic      | 200     |
| ESXn (M710s)                    | vSwitch0         | vmnic0, vmnic1, vmnic2, vmnic3  | Management       | 100     |
|                                 |                  |                                 | LAN Traffic      | 200     |
|                                 | vSwitch1         | vmnic8 (vmnic4, vmnic9, vmnic5) | iSCSI Traffic #1 | NONE    |
|                                 |                  | vmnic4 (vmnic8, vmnic5, vmnic9) | iSCSI Traffic #2 |         |
|                                 |                  | vmnic9 (vmnic5, vmnic8, vmnic4) | iSCSI Traffic #3 |         |
| vmnic5 (vmnic9, vmnic4, vmnic8) | iSCSI Traffic #4 |                                 |                  |         |

A detailed view of one Port Group of the ‘iSCSI Traffic’ virtual switch, reporting the network adapters relationship (active versus standby), is shown in Figure 17.



**Figure 17 Configuration – Port Group details**

## A.4.2 Virtual Machines network configuration

The assignment of the virtual network adapters for each virtual machine was configured as listed below.

- One virtual network adapter to access the 'LAN Traffic' VLAN
- Four virtual network adapters to access the iSCSI network (only for Exchange mailbox server role)
- MPIO (Multi-path I/O) enabled on the iSCSI network adapters, provided by EqualLogic DSM module
- One virtual network adapter to access the respective network in order to manage or monitor it
- Jumbo frames enabled on the virtual network adapters connected to the SAN network (iSCSI)
- TSO (TCP Segmentation Offload) and LRO (Large Receive Offload) enabled on the virtual network adapters connected to the SAN network (iSCSI)

Table 34 summarizes the assignment of virtual network adapter to each virtual switch and network.

**Table 34 Configuration – Virtual Machines network adapters**

| VM      | #vNIC | Adapter Type | Virtual Switch | Port Group       | VLAN ID |
|---------|-------|--------------|----------------|------------------|---------|
| DC01    | #1    | E1000        | vSwitch0       | LAN Traffic      | 200     |
| vCENTER | #1    | E1000        | vSwitch0       | LAN Traffic      | 200     |
|         | #1    | E1000        |                | vCenters         | 100     |
| MBXn    | #1    | E1000        | vSwitch0       | LAN Traffic      | 200     |
|         | #2    | VMXNET3      | vSwitch1       | iSCSI Traffic #1 | NONE    |
|         | #3    | VMXNET3      |                | iSCSI Traffic #2 |         |
|         | #4    | VMXNET3      |                | iSCSI Traffic #3 |         |
|         | #5    | VMXNET3      |                | iSCSI Traffic #4 |         |

## Appendix B Microsoft Jetstress considerations

Microsoft Exchange Server Jetstress 2010 is a simulation tool able to reproduce the database and logs I/O workload of an Exchange mailbox database role server. It is usually used to verify and validate the conformity of a storage subsystem solution before the full Exchange software stack is deployed. Some key elements and consideration about Microsoft Jetstress are:

- Does not require and should not be hosted on a server where Exchange Server is running
- Performs only Exchange storage access and not host processes simulations. It does not contribute in assessing or sizing the Exchange memory and processes footprints
- Is an ESE application requiring access to the ESE dynamic link libraries to perform database access. It takes advantage of the same API used by the full Exchange Server software stack and as such it is a reliable simulation application
- Runs on a single server. When a multiple servers simulation is required, the orchestration of the distributed instances has to be fostered by external management tools
- Requires, and provides, an initialization step to create and populate the database/s that will be used for the subsequent test phases. The database/s should be planned of the same capacity as the one/s planned for the Exchange Server future deployment
- Its Topology layout includes number and size of simulated mailboxes, number and placement of databases and log files, number of database replica copies (simulates only active databases)
- While carrying out a mailbox profile test, it executes a pre-defined mix of insert, delete, replace and commit operations against the database objects during the transactional step, then it performs a full database checksum
- Collects Application and System Event Logs, performance counter values for key performance indicators of both operating system resources and ESE instances during transactional and DB checksum phases. It then generates a detailed HTML-based report
- Throttles the disk I/O generation using the assigned IOPS per mailbox, thread count per database and SluggishSessions threads property (fine tuning for threads execution pace)

For additional information about Exchange Jetstress 2010 refer to Microsoft documentation:  
*Tools for performance and Scalability Evaluation*, available at:  
<http://technet.microsoft.com/en-us/library/dd335108.aspx>

## Additional resources

The following Dell publications are referenced in this document or are recommended sources for additional information.

- Dell EqualLogic Configuration Guide  
<http://www.delltechcenter.com/page/EqualLogic+Configuration+Guide>
- Dell EqualLogic PS Series Group Administration Guide  
[https://support.equallogic.com/support/download\\_file.aspx?id=1125](https://support.equallogic.com/support/download_file.aspx?id=1125)
- Dell PowerEdge Blade Server and Enclosure Documentation  
<http://support.dell.com/support/edocs/systems/pem/en/index.htm>
- Dell PowerConnect 70xx Documentation  
<http://support.dell.com/support/edocs/network/PC70xx/en/index.htm>
- Dell PowerConnect M6220 Documentation  
<http://support.dell.com/support/edocs/network/PCM6220/en/index.htm>
- Dell PowerConnect M6348 Documentation  
<http://support.dell.com/support/edocs/NETWORK/PCM6348/en/index.htm>
- Sizing and Best Practices for Microsoft Exchange 2010 on VMware vSphere and EqualLogic storage  
<http://www.delltechcenter.com/page/Sizing+and+Best++Practices+for+Microsoft+Exchange+2010+on+VMware+vSphere+and+EqualLogic+Storage>
- Best Practices for Enhancing Microsoft Exchange Server 2010 Data Protection and Availability using Dell EqualLogic Snapshots  
<http://www.delltechcenter.com/page/Enhancing+Microsoft+Exchange+Server+2010+Data+P+rotection+and+Availability+with+EqualLogic+Snapshots>

The following Microsoft publications are referenced in this document or are recommended sources for additional information.

- Understanding the Exchange 2010 Store  
<http://technet.microsoft.com/en-us/library/bb331958.aspx>
- Understanding Database Availability Groups  
<http://technet.microsoft.com/en-us/library/dd979799.aspx>
- Understanding the Mailbox Database Cache  
<http://technet.microsoft.com/en-us/library/ee832793.aspx>
- Tools for performance and Scalability Evaluation  
<http://technet.microsoft.com/en-us/library/dd335108.aspx>
- Understanding Mailbox Database and Log Capacity Factors  
<http://technet.microsoft.com/en-us/library/ee832796.aspx>
- Supported and tested Microsoft iSCSI Software Target limits  
<http://support.microsoft.com/kb/2535811>



THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.