Dell EMC PowerEdge R740xd 10,000 Mailbox Resiliency Microsoft Exchange 2016 Storage Solution

Tested with ESRP – Storage Version 4.0
Tested Date: August 2017

Table of Contents

1 Overview

This document provides information about Dell EMC's storage solution for Microsoft Exchange Server. This solution is based on the Microsoft Exchange Solution Reviewed Program (ESRP) – Storage program v4.0. For any questions or comments regarding the contents of this document, see Additional Information.

The ESRP – Storage program was developed by Microsoft Corporation to provide a common storage testing framework for vendors to provide storage solutions for Microsoft Exchange Server. For more information about the Microsoft ESRP — Storage program, see https://technet.microsoft.com/en-us/office/dn756396.aspx

This technical white paper discusses Dell EMC's solution for 10,000 Exchange mailboxes with 20GB mailbox size supporting up to 150 messages per day in a four-copy DAG. The solution uses the Dell EMC PowerEdge R740xd server for the Exchange mailbox server role and uses the internal storage of PowerEdge R740xd for storing the Exchange mailbox databases and transaction logs.

1.1 Disclaimer

This document has been produced independently of Microsoft Corporation. Microsoft Corporation expressly disclaims responsibility for, and makes no warranty, express or implied, with respect to the accuracy of the contents of this document.

The information in this document represents the current view of Dell EMC on the issues discussed as of the date of publication. Due to changing market conditions, it should not be interpreted to be a commitment on the part of Dell EMC and Dell EMC cannot guarantee the accuracy of any information presented after the date of publication.

2 Features

This technical white paper describes a tested and validated storage solution for a 10,000 mailbox Exchange 2016 site-resilient environment by using Database Availability Group (DAG). A DAG is a high-availability (HA) mechanism in Microsoft Exchange 2016 that supports multiple copies (up to 16) of Exchange database. There can be only one active copy of a given Exchange 2016 database at any given time. Mail clients access the active copy, and changes to the active copy are replicated with database copies on other members of the same DAG. All the servers within a DAG are configured to be identical in terms of storage resources for Exchange 2016 databases and logs. The active and passive copies do not share any storage resources, as they reside on their own dedicated storage resources.

Microsoft preferred architecture for Exchange 2016 recommends having an Active/Active multi-site deployment with four database copies equally distributed between both sites. This mailbox resiliency solution includes a single DAG and four copies of each mailbox database, spanning two sites: Site A and Site B. Both sites have active databases hosting 5000 mailboxes each--together 10,000 mailboxes during normal operation. The tested environment simulates up to 10,000 users with 20GB Mailbox size and 150 messages a day or 0.121 IO operations per second (IOPS) per user, including a 20% IO headroom.

In this solution, the PowerEdge R740xd server with 3.5-inch drives is configured for the Mailbox Server role. The 3.5-inch chassis of PowerEdge R740xd server has a distinct configuration mode, where twelve 3.5-inch drives can be placed in the front of the chassis (front bays) and four 3.5-inch drives can be placed in the internal hard-drive tray of the chassis (mid bay). In addition to this, rear bays can accommodate up to four 2.5-inch drives or up to two 3.5-inch drives. Thus, the PowerEdge R740xd server provides extra storage compared to the PowerEdge R730xd server. Number of active databases hosted on one PowerEdge R740xd server during normal runtime is 14. In the event of a server failure in one of the sites, the surviving PowerEdge R740xd servers in the site host 16 active databases.

In the event of a site failure, each PowerEdge R740xd server in the available site hosts 28 active databases. In case of one server failure in the only available site, each server hosts 32 active databases. Following are the major features of the server/storage system:

- Capable of hosting up to sixteen 3.5-inch Large Form Factor (LFF) SAS/Nearline (NL) SAS/SATA drives of up to 10 TB including the four drives in the mid bay of the chassis, plus four additional 2.5-inch disk drives in the back of the system (the 3.5-inch LFF configuration of the PowerEdge R740xd is used as part of this solution); or up to twenty-eight 2.5-inch Small Form Factor (SFF) SAS/NL SAS/SATA drives of up to 2 TB[1] capacity (including the four 2.5-inch back-accessible disk drives) or up to eighteen 1.8-inch hard drives of up to 960GB in addition to eight 3.5-inch Large Form Factor drives.
- Host-based RAID options with Dell EMC PowerEdge RAID Controllers--PERC H730p, H740p, HBA330, and Software RAID (SWRAID) S140
- Host-based RAID options with Dell EMC PowerEdge RAID Controller H840 for external drives

[1] This information is accurate as of the date written.

3 Solution Components

The solution employs building blocks consisting of Dell EMC PowerEdge R740xd servers that are capable of meeting the high performance requirements of messaging deployments. The solution is for up to 10,000 mailboxes with 20GB mailbox size. The following subsections describe the hardware components that are part of this Exchange solution:

Figure 1 Dell EMC PowerEdge R740xd 3.5-inch server

3.1 Dell EMC PowerEdge R740xd

Dell EMC PowerEdge R740xd is a 2-socket, 2U, rack server with a highly expandable memory, dense storage capacity and impressive I/O capabilities. The PowerEdge R740xd server can readily handle data-intensive applications that require large storage capacity and I/O performance, such as email. It delivers the performance and availability required for mission-critical email services and is a great hardware building block for midsize to large organizations.

The internal RAID controller provides a range of RAID levels for improved storage reliability, while the optional CacheCade feature caches the most frequently accessed data, thus boosting database performance. The major features of the server or storage system are as follows. The following simply gives the variety and range of what R740xd can offer. For Exchange Sever, we have selected specific configurations to optimize performance and follow the best pratices guidance given by Microsoft. Please see details of such configurations in Section 4.3 of this paper.

- Up to two Intel® Xeon® Scalable processors with up to 28 cores per processor
- 24 DDR4 DIMM slots that support RDIMM /LRDIMM, speeds up to 2666MHz, 3TB max. Up to 12 NVDIMM, 192 GB Max.
- Up to 160TB Maximum Raw Internal Storage within front and mid bay.
- Choice of chassis configuration with sixteen 3.5-inch LFF disk drives, twenty-four 2.5-inch SFF disk drives, or eighteen 1.8-inch disk drives along with eight 3.5-inch LFF disk drives
- Front loading drive bays plus four 2.5-inch SFF back-accessible drives
- Integrated RAID support through PERC H730p, H740p, HBA330, Software RAID (SWRAID) S140 and External RAID support through PERC H840
- Riser options with up to eight PCIe 3.0 expansion slots
- Choice of NIC technologies
- Dell EMC OpenManage portfolio of systems management solutions, including:
    - OpenManage Essentials console
    - iDRAC9 with Lifecycle Controller

The PowerEdge R740xd chassis configured with the 3.5-inch large form factor drives is used as part of this solution. For more information, see Dell EMC PowerEdge R740xd Server product page.

3.2 PowerEdge RAID Controller H740P

The Dell EMC PERC (PowerEdge RAID Controller) family of enterprise-class controllers is designed for enhanced performance, increased reliability and fault tolerance, and simplified management. They provide a powerful, easy-to-manage solution to create a robust infrastructure and help maximize server uptime.

PERC H740P is used in the PowerEdge R740xd server that hosts the Exchange Server. PERC H740P supports 6Gb/s and 12Gb/s SAS or SATA hard-disk drives and solid-state drives. PERC H740P is the internal host-based RAID Controller card from the PERC Series 10 family. These PERC cards, built on the LSI SAS 3508 dual-core ARM A15 Processor RAID-on-Chip (ROC), offer unmatched I/O performance for databases, applications and streaming digital media environments.

Table 1 shows the technical specifications of PERC H740P. For more information, see Dell EMC PowerEdge RAID Controller product page.

Table 1 Dell EMC PowerEdge RAID controller H740P technical specifications

Feature	**Specification**
Solution provided	Eight-port external SAS solution for performance-hungry external storage environments
Form factor	PCIe Adapter Card and Mini Monolithic
Connectors	Two internal HD Mini-SAS SFF8643
Device support	32 SAS/SATA Devices (PowerEdge Server Largest Drive Configuration)
Host bus type	8-lane, PCI Express 3.1 compliant
Data transfer rate	Up to 12Gbp/s per port
SAS controller	LSISAS 3508 Dual Core ARM A15 Processor - ROC (RAID-On-Chip)
Cache memory	1MB Shared L2 Cache. 6MB On Chip Memory
RAID management	Dell EMC OpenManage Storage Services iDRAC 9 PERC CLI Additional management: • UEFI (HII)

	• CEM
Operating temperature	Maximum ambient temperature: 60°C
Key RAID and data protection features	RAID levels 0, 1, 5, 6 RAID spans 10, 50, 60 Online Capacity Expansion (OCE) Online RAID Level Migration (RLM) Auto resume after power loss during array rebuild or reconstruction/RLM Consistency Check for background data integrity Physical disk power management (Dimmer Switch™) 4K native sector support NVRAM “Wipe” feature protects proprietary data once card is decommissioned SED drive support Load balancing Fast initialization for quick array setup Configurable stripe size up to 1MB Patrol read for media scanning and repair Up to 64 Virtual Drives DDF compliant Configuration on Disk (COD) S.M.A.R.T. support Global and dedicated hot spare with revertible hot-spare support, automatic rebuild, enclosure affinity, and emergency SATA
Operating voltage	+3.3V, +12V and +3.3V_Aux
Optional SSD optimization	Dell EMC FastPath™ firmware feature: delivers high IOPS performance on SSD arrays
Operating systems	Microsoft® Windows Server® 2012 Microsoft® Windows Server® 2016 Red Hat® Enterprise Linux® 6.5 Red Hat® Enterprise Linux® 7.0 or later SUSE® Linux Enterprise Server 12 Virtualization options: • VMware® 6.0 • VMware® 6.5re® 6.5

For more information about recommended hardware specifications, see Section 4.3.

4 Solution Description

In this solution, the PowerEdge R740xd server with 3.5-inch LFF drives is used as the Mailbox Server. PowerEdge R740xd server provides SAS-based internal storage with RAID. The solution uses sixteen 3.5-inch LFF 7.2KRPM NL-SAS disks and four back-accessible 2.5-inch disk drives in the PowerEdge R740xd server in the following layout:

- Two back-accessible disk drives (in RAID 1 container) for the operating system plus application files
- Two back-accessible disk drives (in RAID 1 container) for the Exchange Transport database
- Fourteen disk drives (in RAID 0 containers) for the Exchange database and its transaction logs
- One disk drive marked for Restore LUN
- One disk drive marked for Auto Reseed volume

The solution has a 4-copy DAG Layout (Active/Active) with Exchange servers distributed between two sites: Site A and Site B. Each server node has 14-RAID 0 LUNs hosting one active and three passive databases per LUN during normal operating conditions. Each of these databases hosts 45 users with 20GB mailbox size per user. Thus, a single server can accommodate 625 mailboxes in normal runtime. Sixteen such servers (8 each in Site A and Site B) provide Exchange mailbox services for 10,000 users as outlined in this white paper. In the event of a server failure in one of the sites, corresponding passive database copies in the surviving servers get activated. The design has also considered the site failure and the first server failure in the surviving site, which is the worst case scenario. In this case, seven surviving servers in the available site can accommodate 1,429 mailboxes across 32 databases, and this scenario has been tested.

Active Database Activation Configuration (DC1 Site Failure) / DAG	/ Datacenter 2 Active Server	Active Mailboxes / Server	/ Datacenter 2	Total Active Mailboxes in DC2
Number of Active Databases (Secondary Datacenter (DC2) Activation)	28	1250	224	10000
Number of Active Databases (First Server Failure after SDC Activation)	32	1429	224	10000

The mailbox user profile that was tested had 150 messages per day or 0.121 IOPS per user, which included a 20% IO overhead.

Figure 2 and Figure 3 represent the distribution of database copies across the DAG members in Site A and Site B. It shows a 4-copy Active/Active DAG site resilient solution with Exchange Servers hosted at both sites.

Figure 2 Database Availability Group architectural diagram – Site A

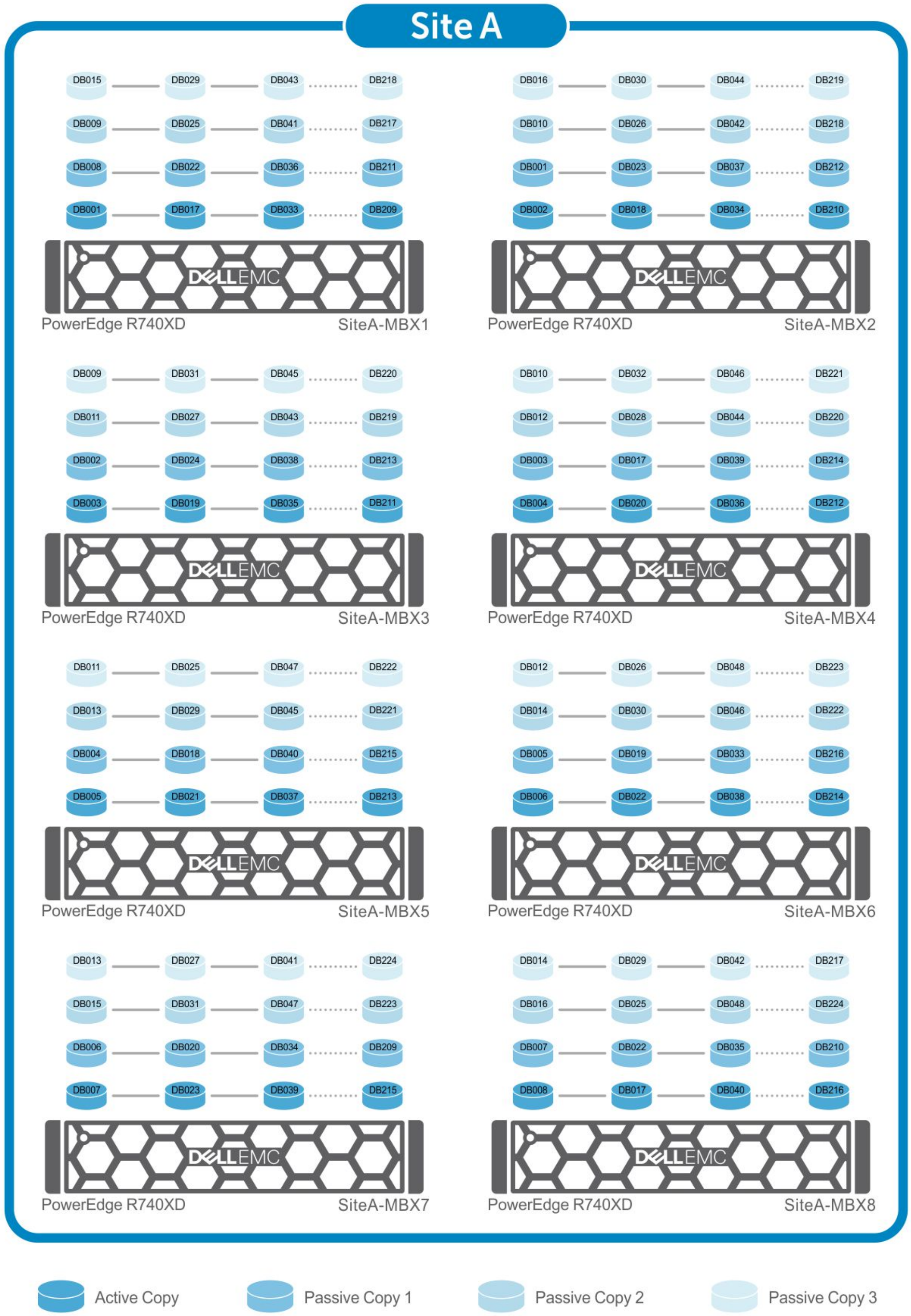
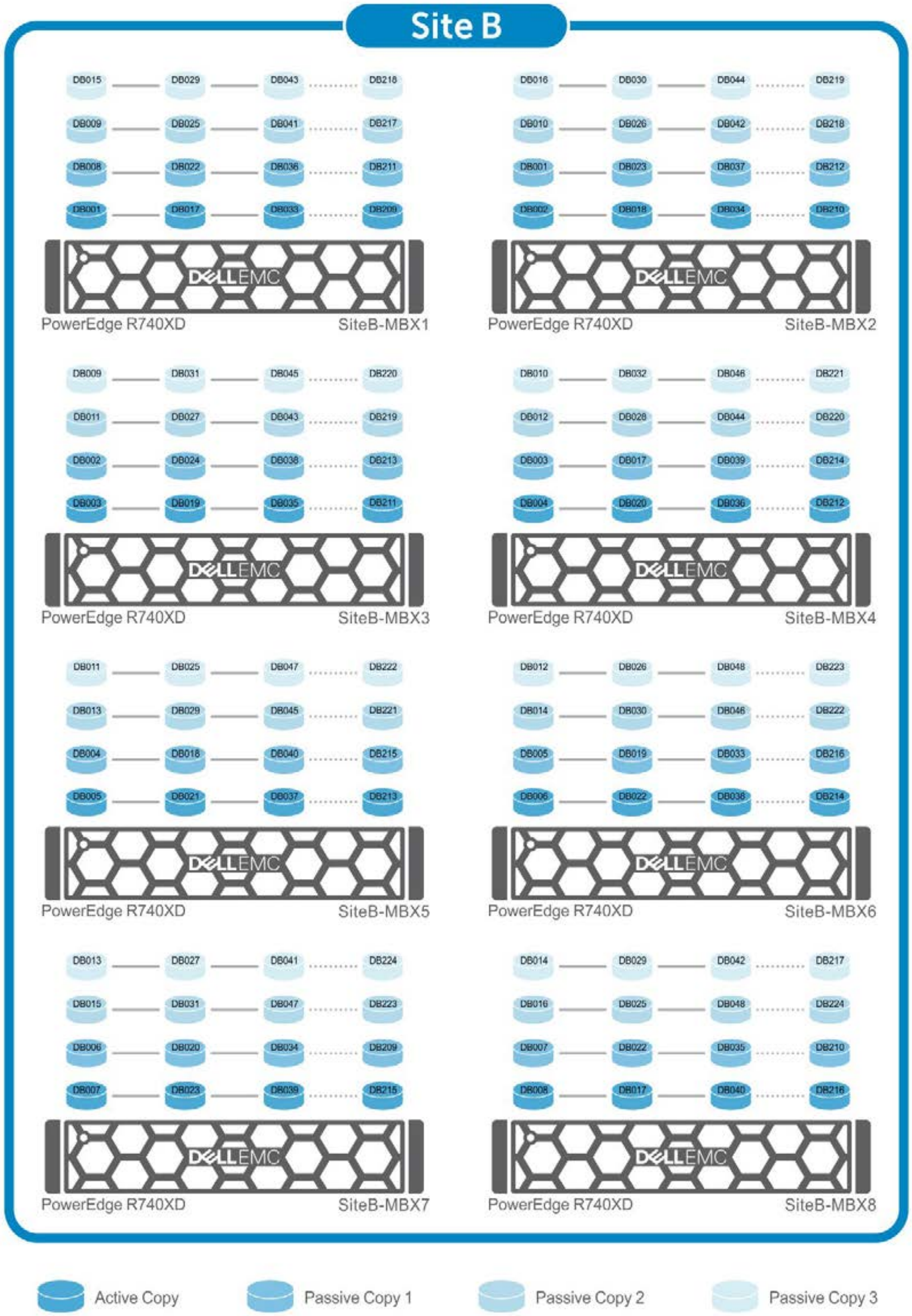


Figure 3 Database Availability Group architectural diagram – Site B



4.1 Failure and Recovery Scenarios

Figure 2 and Figure 3 show the logical diagram of the solution at Site A and Site B, where each site hosts eight servers. A single server failure in Site A or Site B activates the passive copies of the impacted databases. If there is a complete site failure, then the passive copies in the surviving site get activated, and the users connect to their databases on the surviving site. When one site is completely unavailable, and one of the servers in the surviving site fails, the corresponding passive database copies in the surviving servers get activated as shown in Figure 4. This condition is simulated in the test and considered the worst-case failure. Thus, each server is designed in a way that any one server is capable of holding the additional load. Each server is capable of handling the load for 1,429 mailboxes. Therefore, with seven servers, all 10,000 mailboxes can be managed without compromising on the performance.

Figure 4 Worst-case failure scenario

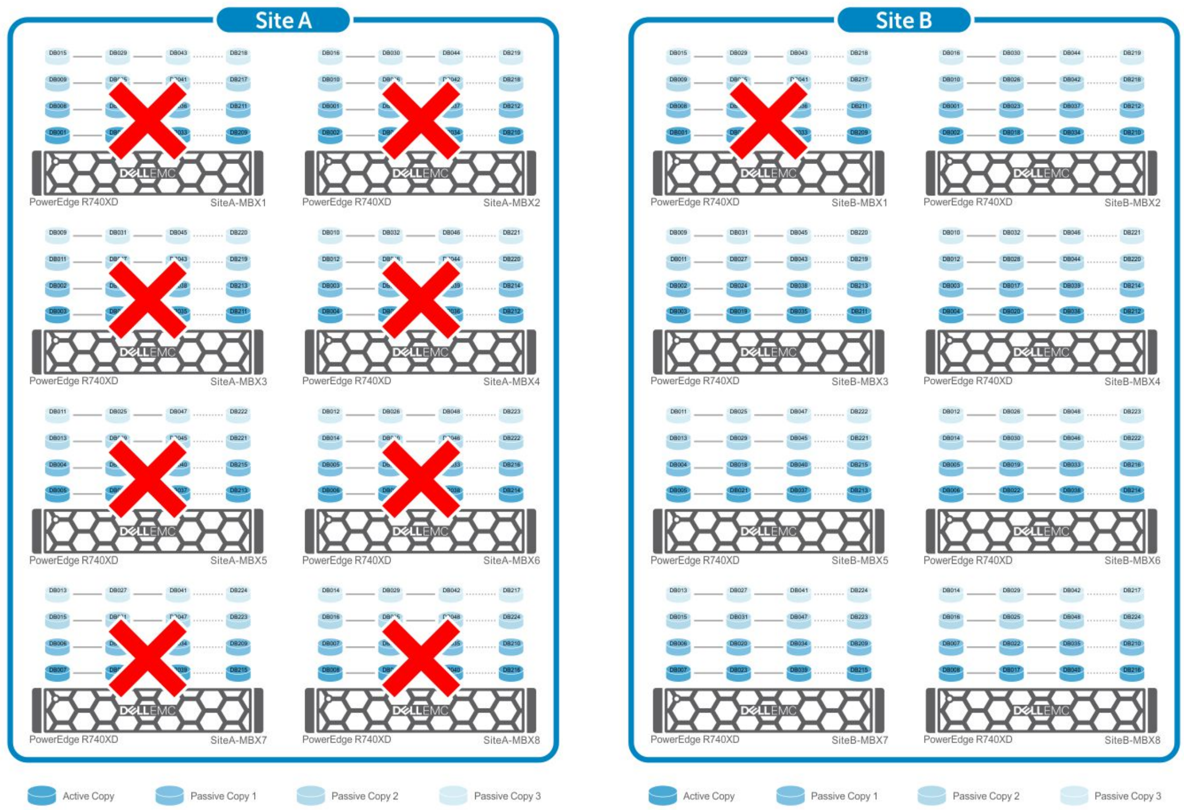


Figure 5 represents the database distribution across servers. [Microsoft Exchange 2013 Server Role Requirements Calculator]() can be used to derive the database distribution including the active and passive copies across servers located in both Site A and Site B. The database distribution follows a particular pattern to ensure that if a server fails, the passive copies are activated on the remaining servers and the load on each server is evenly distributed.

Figure 5 Database/transaction log layout across servers in DAG

Database Name	Active Server	SiteA-MBX1	SiteA-MBX2	SiteA-MBX3	SiteA-MBX4	SiteA-MBX5	SiteA-MBX6	SiteA-MBX7	SiteA-MBX8	SiteB-MBX1	SiteB-MBX2	SiteB-MBX3	SiteB-MBX4	SiteB-MBX5	SiteB-MBX6	SiteB-MBX7	SiteB-MBX8
DB001	SiteA-MBX1	1	2							3	4						
DB002	SiteA-MBX2		1	2							3	4					
DB003	SiteA-MBX3			1	2							3	4				
DB004	SiteA-MBX4				1	2							3	4			
DB005	SiteA-MBX5					1	2							3	4		
DB006	SiteA-MBX6						1	2							3	4	
DB007	SiteA-MBX7							1	2							3	4
DB008	SiteA-MBX8	2							1	4							3
DB009	SiteB-MBX1	3		4						1		2					
DB010	SiteB-MBX2		3		4						1		2				
DB011	SiteB-MBX3			3		4						1		2			
DB012	SiteB-MBX4				3		4						1		2		
DB013	SiteB-MBX5					3		4						1		2	
DB014	SiteB-MBX6						3		4						1		2
DB015	SiteB-MBX7	4						3		2						1	
DB016	SiteB-MBX8		4						3		2						1
DB017	SiteA-MBX1	1			2					3			4				
DB018	SiteA-MBX2		1			2					3			4			
DB019	SiteA-MBX3			1			2					3			4		
DB020	SiteA-MBX4				1			2					3			4	
DB021	SiteA-MBX5					1			2					3			4
DB022	SiteA-MBX6	2					1			4					3		
DB023	SiteA-MBX7		2					1			4					3	
DB024	SiteA-MBX8			2					1			4					3
DB025	SiteB-MBX1	3				4				1				2			
DB026	SiteB-MBX2		3				4				1				2		
DB027	SiteB-MBX3			3				4				1				2	
DB028	SiteB-MBX4				3				4				1				2
DB029	SiteB-MBX5	4				3				2				1			
DB030	SiteB-MBX6		4				3				2				1		
DB031	SiteB-MBX7			4				3				2				1	
DB032	SiteB-MBX8				4				3				2				1
DB033	SiteA-MBX1	1					2			3					4		
DB034	SiteA-MBX2		1					2			3					4	
DB035	SiteA-MBX3			1					2			3					4
DB036	SiteA-MBX4	2			1					4			3				
DB037	SiteA-MBX5		2			1					4			3			
DB038	SiteA-MBX6			2			1					4			3		
DB039	SiteA-MBX7				2			1					4			3	
DB040	SiteA-MBX8					2			1					4			3
DB041	SiteB-MBX1	3						4		1						2	
DB042	SiteB-MBX2		3						4		1						2
DB043	SiteB-MBX3	4		3						2		1					
DB044	SiteB-MBX4		4		3						2		1				
DB045	SiteB-MBX5			4		3						2		1			
DB046	SiteB-MBX6				4		3						2		1		
DB047	SiteB-MBX7					4		3						2		1	
DB048	SiteB-MBX8						4		3						2		1
DB049	SiteA-MBX1	1							2	3							4

DB050	SiteA-MBX2	2	1							4	3						
DB051	SiteA-MBX3		2	1							4	3					
DB052	SiteA-MBX4			2	1							4	3				
DB053	SiteA-MBX5				2	1							4	3			
DB054	SiteA-MBX6					2	1							4	3		
DB055	SiteA-MBX7						2	1							4	3	
DB056	SiteA-MBX8							2	1							4	3
DB057	SiteB-MBX1	3	4							1	2						
DB058	SiteB-MBX2		3	4							1	2					
DB059	SiteB-MBX3			3	4							1	2				
DB060	SiteB-MBX4				3	4							1	2			
DB061	SiteB-MBX5					3	4							1	2		
DB062	SiteB-MBX6						3	4							1	2	
DB063	SiteB-MBX7							3	4							1	2
DB064	SiteB-MBX8	4							3	2							1
DB065	SiteA-MBX1	1		2						3		4					
DB066	SiteA-MBX2		1		2						3		4				
DB067	SiteA-MBX3			1		2						3		4			
DB068	SiteA-MBX4				1		2						3		4		
DB069	SiteA-MBX5					1		2						3		4	
DB070	SiteA-MBX6						1		2						3		4
DB071	SiteA-MBX7	2						1		4						3	
DB072	SiteA-MBX8		2						1		4						3
DB073	SiteB-MBX1	3			4					1			2				
DB074	SiteB-MBX2		3			4					1			2			
DB075	SiteB-MBX3			3			4					1			2		
DB076	SiteB-MBX4				3			4					1			2	
DB077	SiteB-MBX5					3			4					1			2
DB078	SiteB-MBX6	4					3			2					1		
DB079	SiteB-MBX7		4					3			2					1	
DB080	SiteB-MBX8			4					3			2					1
DB081	SiteA-MBX1	1				2				3				4			
DB082	SiteA-MBX2		1				2				3				4		
DB083	SiteA-MBX3			1				2				3				4	
DB084	SiteA-MBX4				1				2				3				4
DB085	SiteA-MBX5	2				1				4				3			
DB086	SiteA-MBX6		2				1				4				3		
DB087	SiteA-MBX7			2				1				4				3	
DB088	SiteA-MBX8				2				1				4				3
DB089	SiteB-MBX1	3					4			1					2		
DB090	SiteB-MBX2		3					4			1					2	
DB091	SiteB-MBX3			3					4			1					2
DB092	SiteB-MBX4	4			3					2			1				
DB093	SiteB-MBX5		4			3					2			1			
DB094	SiteB-MBX6			4			3					2			1		
DB095	SiteB-MBX7				4			3					2			1	
DB096	SiteB-MBX8					4			3					2			1
DB097	SiteA-MBX1	1						2		3						4	
DB098	SiteA-MBX2		1						2		3						4
DB099	SiteA-MBX3	2		1						4		3					
DB100	SiteA-MBX4		2		1						4		3				
DB101	SiteA-MBX5			2		1						4		3			
DB102	SiteA-MBX6				2		1						4		3		

DB103	SiteA-MBX7					2		1						4		3	
DB104	SiteA-MBX8						2		1						4		3
DB105	SiteB-MBX1	3							4	1							2
DB106	SiteB-MBX2	4	3							2	1						
DB107	SiteB-MBX3		4	3							2	1					
DB108	SiteB-MBX4			4	3							2	1				
DB109	SiteB-MBX5				4	3							2	1			
DB110	SiteB-MBX6					4	3							2	1		
DB111	SiteB-MBX7						4	3							2	1	
DB112	SiteB-MBX8							4	3							2	1
DB113	SiteA-MBX1	1	2							3	4						
DB114	SiteA-MBX2		1	2							3	4					
DB115	SiteA-MBX3			1	2							3	4				
DB116	SiteA-MBX4				1	2							3	4			
DB117	SiteA-MBX5					1	2							3	4		
DB118	SiteA-MBX6						1	2							3	4	
DB119	SiteA-MBX7							1	2							3	4
DB120	SiteA-MBX8	2							1	4							3
DB121	SiteB-MBX1	3		4						1		2					
DB122	SiteB-MBX2		3		4						1		2				
DB123	SiteB-MBX3			3		4						1		2			
DB124	SiteB-MBX4				3		4						1		2		
DB125	SiteB-MBX5					3		4						1		2	
DB126	SiteB-MBX6						3		4						1		2
DB127	SiteB-MBX7	4						3		2						1	
DB128	SiteB-MBX8		4						3		2						1
DB129	SiteA-MBX1	1			2					3			4				
DB130	SiteA-MBX2		1			2					3			4			
DB131	SiteA-MBX3			1			2					3			4		
DB132	SiteA-MBX4				1			2					3			4	
DB133	SiteA-MBX5					1			2					3			4
DB134	SiteA-MBX6	2					1			4					3		
DB135	SiteA-MBX7		2					1			4					3	
DB136	SiteA-MBX8			2					1			4					3
DB137	SiteB-MBX1	3				4				1				2			
DB138	SiteB-MBX2		3				4				1				2		
DB139	SiteB-MBX3			3				4				1				2	
DB140	SiteB-MBX4				3				4				1				2
DB141	SiteB-MBX5	4				3				2				1			
DB142	SiteB-MBX6		4				3				2				1		
DB143	SiteB-MBX7			4				3				2				1	
DB144	SiteB-MBX8				4				3				2				1
DB145	SiteA-MBX1	1					2			3					4		
DB146	SiteA-MBX2		1					2			3					4	
DB147	SiteA-MBX3			1					2			3					4
DB148	SiteA-MBX4	2			1					4			3				
DB149	SiteA-MBX5		2			1					4			3			
DB150	SiteA-MBX6			2			1					4			3		
DB151	SiteA-MBX7				2			1					4			3	
DB152	SiteA-MBX8					2			1					4			3
DB153	SiteB-MBX1	3						4		1						2	
DB154	SiteB-MBX2		3						4		1						2
DB155	SiteB-MBX3	4		3						2		1					
DB156	SiteB-MBX4		4		3						2		1				

DB157	SiteB-MBX5			4		3						2		1			
DB158	SiteB-MBX6				4		3						2		1		
DB159	SiteB-MBX7					4		3						2		1	
DB160	SiteB-MBX8						4		3						2		1
DB161	SiteA-MBX1	1							2	3							4
DB162	SiteA-MBX2	2	1							4	3						
DB163	SiteA-MBX3		2	1							4	3					
DB164	SiteA-MBX4			2	1							4	3				
DB165	SiteA-MBX5				2	1							4	3			
DB166	SiteA-MBX6					2	1							4	3		
DB167	SiteA-MBX7						2	1							4	3	
DB168	SiteA-MBX8							2	1							4	3
DB169	SiteB-MBX1	3	4							1	2						
DB170	SiteB-MBX2		3	4							1	2					
DB171	SiteB-MBX3			3	4							1	2				
DB172	SiteB-MBX4				3	4							1	2			
DB173	SiteB-MBX5					3	4							1	2		
DB174	SiteB-MBX6						3	4							1	2	
DB175	SiteB-MBX7							3	4							1	2
DB176	SiteB-MBX8	4							3	2							1
DB177	SiteA-MBX1	1		2						3		4					
DB178	SiteA-MBX2		1		2						3		4				
DB179	SiteA-MBX3			1		2						3		4			
DB180	SiteA-MBX4				1		2						3		4		
DB181	SiteA-MBX5					1		2						3		4	
DB182	SiteA-MBX6						1		2						3		4
DB183	SiteA-MBX7	2						1		4						3	
DB184	SiteA-MBX8		2						1		4						3
DB185	SiteB-MBX1	3			4					1			2				
DB186	SiteB-MBX2		3			4					1			2			
DB187	SiteB-MBX3			3			4					1			2		
DB188	SiteB-MBX4				3			4					1			2	
DB189	SiteB-MBX5					3			4					1			2
DB190	SiteB-MBX6	4					3			2					1		
DB191	SiteB-MBX7		4					3			2					1	
DB192	SiteB-MBX8			4					3			2					1
DB193	SiteA-MBX1	1				2				3				4			
DB194	SiteA-MBX2		1				2				3				4		
DB195	SiteA-MBX3			1				2				3				4	
DB196	SiteA-MBX4				1				2				3				4
DB197	SiteA-MBX5	2				1				4				3			
DB198	SiteA-MBX6		2				1				4				3		
DB199	SiteA-MBX7			2				1				4				3	
DB200	SiteA-MBX8				2				1				4				3
DB201	SiteB-MBX1	3					4			1					2		
DB202	SiteB-MBX2		3					4			1					2	
DB203	SiteB-MBX3			3					4			1					2
DB204	SiteB-MBX4	4			3					2			1				
DB205	SiteB-MBX5		4			3					2			1			
DB206	SiteB-MBX6			4			3					2			1		
DB207	SiteB-MBX7				4			3					2			1	
DB208	SiteB-MBX8					4			3					2			1
DB209	SiteA-MBX1	1						2		3						4	
DB210	SiteA-MBX2		1						2		3						4

DB211	SiteA-MBX3	2		1						4		3					
DB212	SiteA-MBX4		2		1						4		3				
DB213	SiteA-MBX5			2		1						4		3			
DB214	SiteA-MBX6				2		1						4		3		
DB215	SiteA-MBX7					2		1						4		3	
DB216	SiteA-MBX8						2		1						4		3
DB217	SiteB-MBX1	3							4	1							2
DB218	SiteB-MBX2	4	3							2	1						
DB219	SiteB-MBX3		4	3							2	1					
DB220	SiteB-MBX4			4	3							2	1				
DB221	SiteB-MBX5				4	3							2	1			
DB222	SiteB-MBX6					4	3							2	1		
DB223	SiteB-MBX7						4	3							2	1	
DB224	SiteB-MBX8							4	3							2	1

4.2 Storage Sizing

Selecting the right storage is crucial in achieving a balance between cost and performance. The storage size and design should be based on the type of RAID, type of disk drives and number of disk drives—both from capacity and IOPS perspective. The storage design also depends on the actual size of mailbox on the disk drive, the content indexing space, and the required log space.

Microsoft Exchange 2013 Server Role Requirements Calculator can be used to derive the required IOPS for a particular user profile. Figure 6 shows the Mailbox Calculator output for 10,000 users with 150 messages per day profile. The recommended IOPS per server is 302. Microsoft Exchange Jetstress tool verifies if the storage subsystem meets the targeted IOPS requirement. For more information see Section 5.

> **Note:** To calculate the processor, memory and storage sizing for a specific number and size of mailboxes and profiles, you can use the latest version of Exchange Server Role Requirements Calculator published by Microsoft which can be used with both Exchange Server 2013 and Exchange Server 2016

Figure 6 Recommended IOPS from the Microsoft Exchange 2013 server role requirements calculator

Host IO and Throughput Requirements	/ Database	/ Server	/ DAG	/ Environment
Total Database Required IOPS	5	302	4824	4824
Total Log Required IOPS	1	64	1020	1020
Database Read I/O Percentage	60%	--	--	--
Background Database Maintenance Throughput Requirements	1.0 MB/s	56 MB/s	896 MB/s	896 MB/s

4.3 Recommended Hardware Configuration

Table 2 and Table 3 provide the server and storage configuration as well as the driver and firmware versions used in the tested solution.

Table 2 Exchange Server configuration

Microsoft Exchange Server System	**Dell EMC PowerEdge R740xd Server with 3.5” HDD Chassis**
CPU	2 × Intel Xeon Gold 5115 processor @ 2.40GHz with 10-cores
Memory	Up to 96GB DDR4[2]
NIC	Broadcom NetXtreme Gigabit Ethernet
RAID Controller	PERC H740P Adapter Firmware version: 50.0.1-0537 Storport Driver Version 10.0.14393.351 Driver version 7.700.51.00
Internal Disks	4 x 1.2TB SAS 2.5-inch 10K RPM disk drives • 2 x 1.2TB SAS 2.5-inch 10K RPM disk drives in RAID 1 volume (Operating System and Application) • 2 x 1.2TB SAS 2.5-inch 10K RPM disk drives in RAID 1 volume (Exchange queue database)

[2] Microsoft recently raised the maximum memory to 192GB to accommodate the hardware availability in terms of 96GB memory support; but that doesn’t change the fact that 96GB memory is still the right threshold/standard for our R740xd configuration used in this solution for the JetStress test.

Table 3 Storage subsystem configuration (internal storage)

Storage Subsystem	**Dell EMC PowerEdge R740xd Internal 3.5-inch drives**
Disks	16 x 8TB 7.2K RPM NL-SAS 3.5-inch disk: • 14 x 8TB 7.2K RPM NL-SAS 3.5-inch drive in 14 x RAID 0 volumes (for DB and Log) • 1 x 8TB 7.2K RPM NL-SAS 3.5-inch drive (for Restore LUN) • 1 x 8TB 7.2K RPM NL-SAS 3.5-inch drive (Auto Reseed Volume)
RAID Controller	Dell EMC PowerEdge RAID Controller H740P (Firmware version: 50.0.1-0537)

5 Targeted Customer Profile

This solution is intended for midsize to large organizations hosting up to 10,000 Exchange 2016 mailboxes. The configuration used for testing was as follows:

- Number of mailboxes: 10,000
- Number of sites: 2 (Site A and Site B)
- Number of servers in each site: 8 in Site A and 8 in Site B
- User IO profile: 150 messages sent and received or 0.121 IOPS per mailbox (This includes 20% IO overhead factor)
- 20GB Mailbox quota per mailbox
- 24x7 Background Database Maintenance enabled
- Database Availability Group (DAG) for Mailbox Resiliency (4 copies simulated-1 Active, 3 Passive)

5.1 Tested User Profile

The tested user profile had 0.121 IOPS per user with a 20GB mailbox size. This equates to 150 messages (sent or received) per mailbox per day and accounts for an additional 20% I/O overhead. Additional applications such as certain mobile messaging applications can increase the IOPS profile of a user by three to four times.

5.2 Tested Deployment

The tested deployment simulated a site level failure scenario where one member of the surviving site (Site A or Site B) was completely unavailable and the passive copies on the surviving DAG members were activated to provide mailbox service continuity. Therefore, the IOPS required for 1,429 mailboxes were simulated on one of the surviving servers. The target IOPS for the given profile was172.9. The achieved IOPS was 797—much higher than the target—and the solution still maintained read and write latencies well within the recommended thresholds. The following tables summarize the testing environment:

Table 4 Simulated Exchange configuration

Feature	Specification
Number of Sites	2 (Site A and Site B)
Number of servers per DAG	16 (8 in Site A and 8 in Site B), 7 servers hosting active mailboxes during the test
Number of active mailboxes per server	625 (during normal operations) 1,250 (during site failure) 1,429 (in tested failover during worst-case failure)

Feature	Specification
Number of databases per server	32
Number of copies per database	4 (2 in Site A and 2 in Site B)
Number of mailboxes per database	45
Simulated profile: IOPS per mailbox	0.121 (150 messages/day) This includes 20% IO overhead factor
Database/Log LUN size	7.27TB
Number of LUNs per server	14 (8 tested)
Number of DBs per LUN	4
Background database maintenance (BDM)	Tested with BDM enabled
Total database size for performance testing	1.45TB per DB 81.2TB total
Percentage of storage capacity used by the Exchange database	81.2TB / 101.78TB 79.78%

Table 5 Storage and server hardware

Feature	Specification
Storage connectivity (Fiber Channel, SAS, SATA, iSCSI)	SAS

Feature	Specification
Storage model and OS/firmware revision	Dell EMC PowerEdge R740xd with PERC H740P attached to internal drives Firmware 50.0.1-0537
Storage cache	8GB non-volatile cache memory
Number of storage controllers	1, attached to internal drives in PowerEdge R740xd
Number of storage ports	2 (Two internal HD Mini-SAS SFF8643)
Maximum bandwidth of storage connectivity to server	12Gb/s per port
Switch type/model/firmware revision	NA
HBA model and firmware	H740P Firmware 50.0.1-0537
Number of HBAs per server	1
Host server type	Dell EMC PowerEdge R740xd 2 x Intel Xeon processor 96GB RAM
Total number of disks tested in the solution	56 (8 per server)
Maximum number of spindles that can be hosted in the storage	16 x 3.5” and 4 x 2.5” per server

Table 6 Storage and server software

Feature	Specification
HBA driver	PERC H740P SAS-RAID 7.700.51.00
HBA QueueTarget Setting	N/A
HBA QueueDepth Setting	N/A
Multipathing	N/A
Host OS	Windows Server 2016 Data Center X64 Edition
ESE.dll file version	15.1.1034.26
Replication solution name/version	N/A

Table 7 Storage disk configuration (Mailbox store disks)

Feature	Specification
Disk type, speed and firmware revision	DELL EMC 7.2K 3.5" RPM 8TB NL-SAS Model – ST8000NM0075
Raw capacity per disk (TB)	7.27TB
Number of physical disks in the test	56 (8 per Server)
Total raw storage capacity (TB)	407TB (58TB per Server)
RAID level	RAID 0
Number of disks per LUN	1
Total formatted capacity	7.27TB per LUN 58.16TB per server
Storage capacity utilization	58.16/407=14.3% Formatted capacity/Total raw capacity
Database capacity utilization	(1.45TBx32)/58.18TB=79.78% Database size/Total formatted capacity

5.3 Best Practices

Exchange Server 2010, 2013 and 2016 overcome the memory limitations of earlier Exchange versions by providing support as a 64-bit application. Exchange Server 2016 and Windows Server 2016 provide an additional 64-bit of application server support and server OS support respectively. This provides about 4TB of addressable memory for kernel mode and user mode applications.

Both the application and kernel have sufficient memory for operations, allowing the Extensible Storage Engine (ESE) in Exchange Server 2016 to utilize more memory to buffer data pages. The result is a reduction in the number of Input/Output (I/O) operations, specifically the read operations required for the disk subsystem. The total number of database disk I/O operations for a given user load depends on the available system memory. For a given load, the total database disk Input/Output operations per second (IOPS) decreases over a period with increase in system memory. This decrease in database IOPS is primarily caused by a decrease in database reads.

While sizing the Exchange Storage subsystem, make sure that there are no I/O bottlenecks from an IOPS and disk latency perspective. The disk subsystem should be capable of supporting both the capacity and I/O throughput demands of the application. The following best practices are recommended to improve the I/O subsystem performance:

- For the Exchange 2016 database, the size of elements within a RAID stripe should be set to 512K for best performance.
- Each server should have two RAID 1 volumes – one to host the Operating System and Exchange binaries and the other to host the Exchange queue database. Rest of the storage can be configured as independent RAID 0 volumes.
- Each disk that houses an Exchange database should be formatted with ReFS (with integrity feature disabled) and the DAG should be configured such that AutoReseed formats the disk with ReFS.
- Average database read latencies (Avg. Disk sec/Read) should not exceed 20ms. Exchange Server 2016 storage latencies are most often related to the number of disk drives available for a given workload. Windows Performance Monitor may be used to monitor Exchange Server 2016 database counters.
- Sharing Exchange 2016 storage resources with other applications may negatively affect the performance of Exchange 2016 deployment. Therefore, sharing the spindles that host the Exchange Database and log with any other application or operating system is not recommended.

5.4 Backup Strategy

To protect email data from potential disasters, having a well designed and implemented backup solution is critical. Depending on environmental requirements, different backup strategies such as backup to tape or LAN/SAN-based backup can be implemented. In this solution, a DAG is used to maintain a passive database copy on a separate storage system. This passive copy of the database may be used to back up tape or disk drive.

The log replay test was used to measure the maximum rate at which the log files can be replayed on the passive copies. This is used to determine the restore times and also the database write throughput that can be achieved during a log recovery.

6 Test Result Summary

This section provides a high-level summary of the test data from Microsoft Exchange Jetstress as part of the ESRP requirements.   It also includes the link to the detailed HTML reports, which are generated by the ESRP testing framework.

6.1 Reliability

Reliability tests are run for 24 hours, and the goal is to verify if the storage can handle a high I/O load for a long period. After the stress test, both log and database files are analyzed for integrity to make sure that there is no database/log corruption.

The following list provides an overview of any errors reported during testing:

- Any errors reported in the saved event log file? No
- Any errors reported during the database and log checksum process? No

6.2 Storage Performance Test Result Report

The storage performance test is designed to evaluate the storage with maximum sustainable Exchange I/O for four hours. The test shows how long it takes for storage to respond to an I/O under load. The data in this section 6.2 is the sum of all the logical disks I/Os and average of all the logical disks I/O latency in the four hour test duration. The achieved IOPS was around 798.

As part of the ESRP framework, the Stress Test was also performed. The duration of the test was 24 hours with a target IOPS of 0.121 per user or 173 IOPS per server. The achieved IOPS was around 783 per server. This was well above the target IOPS. The Stress Test Result Report is provided for reference.

6.2.1 Individual Server Metrics

Table 8 shows the sum of I/Os across Mailbox databases and the average latency across all databases on a per server basis.

Table 8 Individual server metrics

Server 1:

Database I/O	
Target Disk Transfers/sec	173
Database Disks Transfers/sec	772
Database Disks Reads/sec	539
Database Disks Writes/sec	233
Average Database Disk Read Latency (ms)	15.21

Database I/O	
Average Database Disk Write Latency (ms)	0.379
Transaction Log I/O	
Log Disks Writes/sec	56.72
Average Log Disk Write Latency (ms)	0.154

Server 2:

Database I/O	
Target Disk Transfers/sec	173
Database Disks Transfers/sec	789
Database Disks Reads/sec	550
Database Disks Writes/sec	239
Average Database Disk Read Latency (ms)	14.8
Average Database Disk Write Latency (ms)	0.18
Transaction Log I/O	
Log Disks Writes/sec	57.64
Average Log Disk Write Latency (ms)	0.11

Server 3:

Database I/O	
Target Disk Transfers/sec	173
Database Disks Transfers/sec	794
Database Disks Reads/sec	555
Database Disks Writes/sec	239
Average Database Disk Read Latency (ms)	15.08
Average Database Disk Write Latency (ms)	0.185
Transaction Log I/O	
Log Disks Writes/sec	61.19
Average Log Disk Write Latency (ms)	0.12

Server 4:

Database I/O	
Target Disk Transfers/sec	173
Database Disks Transfers/sec	819
Database Disks Reads/sec	571
Database Disks Writes/sec	248
Average Database Disk Read Latency (ms)	15.33
Average Database Disk Write Latency (ms)	0.18
Transaction Log I/O	
Log Disks Writes/sec	64.06
Average Log Disk Write Latency (ms)	0.112

Server 5:

Database I/O	
Target Disk Transfers/sec	173
Database Disks Transfers/sec	801
Database Disks Reads/sec	556
Database Disks Writes/sec	245
Average Database Disk Read Latency (ms)	14.97
Average Database Disk Write Latency (ms)	0.184
Transaction Log I/O	
Log Disks Writes/sec	57.59
Average Log Disk Write Latency (ms)	0.099

Server 6:

Database I/O	
Target Disk Transfers/sec	173
Database Disks Transfers/sec	813

Database I/O	
Database Disks Reads/sec	566
Database Disks Writes/sec	247
Average Database Disk Read Latency (ms)	15.36
Average Database Disk Write Latency (ms)	0.18
Transaction Log I/O	
Log Disks Writes/sec	59.01
Average Log Disk Write Latency (ms)	0.098

Server 7:

Database I/O	
Target Disk Transfers/sec	173
Database Disks Transfers/sec	796
Database Disks Reads/sec	554
Database Disks Writes/sec	242
Average Database Disk Read Latency (ms)	15.35
Average Database Disk Write Latency (ms)	0.181
Transaction Log I/O	
Log Disks Writes/sec	57.52
Average Log Disk Write Latency (ms)	0.1

6.2.2 Aggregate Performance Across Servers/DAGs Metrics

Table 9 shows the aggregated results of I/O and the average latency across servers in the solution.

Table 9 Aggregate performance metrics across all servers

Database I/O	
Database Disks Transfers/sec	5584
Database Disks Reads/sec	3891
Database Disks Writes/sec	1693
Average Database Disk Read Latency (ms)	15.16
Average Database Disk Write Latency (ms)	0.21

Database I/O	
Transaction Log I/O	
Log Disks Writes/sec	413.73
Average Log Disk Write Latency (ms)	0.113

6.3 Database Backup/Recovery Performance

There are two test reports in this section. The database backup test measures the sequential read rate of the database files, and the soft recovery test measures the recovery/replay performance (playing transaction logs in to the database).

6.3.1 Database Backup Test Result Report

The test is to measure the maximum rate at which databases could be backed up through VSS. The following table shows the average rate for a single database file:

Table 10 Database backup test metrics

MB read/sec per database	31.39
MB read/sec total per server	1004.33

6.3.2 Soft Recovery Test Result Report

The test is to measure the maximum rate at which the log files can be played on the passive copies. The following table shows the average rate for 505 log files played in a single storage group. Each log file is 1MB in size.

Table 11 Soft recovery test metrics

Average number of log files played	507
Average time to play one Log file (sec)	9.39

7 Detailed Test Results

Detailed Jetstress test results for all seven mailbox servers are attached to this PDF file in the form of HTML reports. Report types are as follows:

- 4-hour performance test
- Checksum for 4-hour performance test
- 24-hour stress test
- Checksum for 24-hour stress test
- Database backup test
- Soft recovery test

7.1 How to View Jetstress Reports

Click the Attachments icon in the left pane of Adobe Reader to view the list of Jetstress reports (HTML files). Double click each item to open the corresponding report in your browser.

8 Conclusion

This ESRP document presents a tested and validated Exchange solution for 10,000 mailboxes with 20GB mailbox size supporting up to 150 messages per day in a four-copy DAG. The solution uses the Dell EMC PowerEdge R740xd server for the Exchange mailbox server role as well as for storing the Exchange mailbox databases and transaction logs.

Testing was carried out as part of the ESRP test framework by using Microsoft Exchange Server 2013 Jetstress which is the recommended tool for Exchange 2013 and Exchange 2016. The test results showed that the proposed solution is more than capable of delivering the IOPS and meeting the capacity requirements to support 10,000 mailboxes with the said mailbox profile.

This document is developed by storage solution providers and reviewed by the Microsoft Exchange Product team. The test results and data presented in this document are based on the tests included in the ESRP test framework. Customers should not quote the data directly for pre-deployment verification. It is necessary to go through the exercises to validate the storage design for a specific customer environment.

The ESRP program is not designed to be a benchmarking program, and the tests are not designed to obtain the maximum throughput for a given solution. Rather, the tests focus on obtaining recommendations from vendors for Exchange application. The data presented in this document should not be used for direct comparisons among solutions.

9 Additional Information

1. **Dell.com/support** is focused on meeting customer requirements with proven services.
2. **DellTechCenter.com** is an IT Community where you can connect with Dell EMC customers and Dell EMC employees to share knowledge, best practices and information about Dell EMC products and installations.
3. Referenced or recommended Dell EMC publications:
    a. Dell EMC IT Consulting
    b. PowerEdge R740xd Rack Server
    c. Dell EMC PowerEdge RAID Controller (PERC) 10 User’s Guide
    d. PowerEdge RAID Controller H740P Data Sheet