Hewlett Packard Enterprise

HPE Reference Architecture for Microsoft Exchange Server 2016 on HPE ProLiant DL380 Gen10 server

Designed for 10,512 users with 21.5GB mailboxes using 10TB large form factor hot plug drives

Reference Architecture

Contents

Executive summary

This Reference Architecture for Microsoft Exchange Server 2016 demonstrates a highly resilient solution that provides large mailboxes for users that expect the utmost in performance and reliability. The solution is based on a building block approach with four servers per site, and 10,512 users per building block, which was tested and validated at 21.5GB mailboxes with a messaging profile of 150 emails sent and received, per user, per day.

This solution is designed using HPE ProLiant DL380 Gen10 servers and Exchange Server 2016, in a building block approach, with a multi-copy database design with Exchange Database Availability Groups (DAGs). This Reference Architecture testing is designed to validate that the storage, CPU and memory subsystems can support this workload in both normal and peak operations, even in a failover scenario, where one of the servers in the primary site is offline. The solution is designed to withstand an outage of either a server within the site or the failover of the entire site and maintain availability to all users. Optimal sizing is done for the failover scenario, such that the CPU load is handled effectively even when the secondary site is unavailable and one of the servers in the primary site is offline, for either planned or unplanned downtime. The solution adequately handles the normal and peak loads for 3,504 users per server with 3 of 4 servers in one site of the building block active.

The solution uses direct-attached disk storage for Exchange databases. Each HPE ProLiant DL380 Gen10 server is configured with 21 total drives; 2 small form factor (SFF) SSDs (800GB) in RAID 1 configuration for the operating system and Exchange Transport files, and 19 large form factor (LFF) hard disk drives, 10TB each, in a RAID-less JBOD configuration. This provides the entire capacity of the 10TB drives for several Exchange databases (4 databases per volume, given a maximum set size of 1.8TB databases). The design and testing shown here in this document cover the importance of understanding how either 2 or 3 of the 4 database copies on a single disk are activated in a failover scenario. It is very important to note that only 2 or 3 copies will be activated on any given disk in the failover scenario where 3 of 4 servers are active and all databases are online.

The test results validate that the Reference Architecture is adequately sized to support 10,512 mailbox users with a 150 message/day profile, with additional performance headroom even in the failover scenario where only three Exchange Servers are online.

All servers were configured identically, except that a processor comparison was designed and delivered. The only difference between the two configurations was that the servers had different processors, as identified below:

- HPE ProLiant DL380 Gen10 server (Intel® Xeon® Gold 6132 processors, 2.6GHz, 14 cores)–run with 12 cores active per socket (24/28 total)
- HPE ProLiant DL380 Gen10 server (Intel Xeon Gold 6126 processors, 2.6GHz, 12 cores)

The HPE servers are configured with an HPE Smart Array P816i-a controller that can operate in mixed mode, which combines RAID and HBA operations simultaneously. However, the HBA mode configuration does not enable controller write caching and therefore should not be used for Exchange databases. This white paper provides the highlights of installing and configuring the server and storage

Target audience: This white paper is intended for messaging architects, Chief Information Officers (CIOs), decision makers, IT support staff, and project managers involved in planning and deploying Microsoft Exchange Server 2016 solutions. A basic level of understanding of Exchange design is helpful, but this document also provides links to Exchange reference material.

Document purpose: The purpose of this document is to showcase performance test results for a Reference Architecture and describe in detail the recommended configuration for this solution, highlighting the distinct advantages of deploying Exchange Server 2016 on the HPE ProLiant DL380 Gen10 system.

This white paper describes testing performed by Hewlett Packard Enterprise in April 2018.

Introduction

This Reference Architecture is built on the features of the HPE ProLiant DL380 Gen10 servers and Exchange Server 2016 for high availability. The HPE ProLiant DL380 Gen10 servers that are part of the solution contain 21 drives; 19 of the 21 drives are large form factor (LFF) hard disk drives, 10TB each. By using what is known as single drive RAID-less JBOD (single drive RAID 0 arrays), we use the entire capacity of the 10TB drives for several Exchange databases (up to 4, given a maximum set size of 1.8TB databases).

This solution is designed using a building block approach, with a multi-copy database design using Exchange Database Availability Groups (DAGs). DAG is the base component of high availability and site resilience built into Exchange. A DAG is a group of up to 16 mailbox servers that host a set of database copies, and thus provide recovery from failures that affect individual servers, databases or internal database corruption.

A DAG provides high availability with multiple copies in both a primary data center location, and a secondary data center or disaster recovery (DR) location. The design here provides two copies of each database in each site. The database copies provide the ability to withstand failures due to either logical corruption in an Exchange database, the failure of one or more disk drives, a single server within the active site, or the complete outage of the servers in the primary data center or site (failover across sites). The user distribution can be either active/passive across the two sites (all databases mounted in the primary or the secondary, not both) or active/active (where user load is distributed between the two sites in any desired proportion).

This design is based on a specific number of servers per DAG which determines the number of users that can be hosted on a single server in a failover scenario. For example, 8 servers (4 servers per site) are sized with 1,314 users per server in normal operations, and 16 servers in the DAG can run up to 2,125 users per server in normal operations (with smaller mailboxes due to the increased user count). The number of servers per DAG is a flexible design decision that customers can make using the tools, resources and design considerations described in this document.

Microsoft has published The Exchange 2016 Preferred Architecture which they describe as the "Exchange Engineering Team's best practice recommendation for what Microsoft believes is the optimum deployment architecture for Exchange 2016, and one that is very similar to what we deploy in Office 365". This HPE Reference Architecture configuration of Exchange follows the Preferred Architecture wherever possible, as it is built using direct-attached storage (DAS) in a RAID-less JBOD configuration. It would also be possible to use RAID 1 mirrors but the mailbox capacity would be reduced to half.

The designs discussed in this document have been validated using the Microsoft Exchange Server Role Requirements Calculator (version 9.1). This solution was modified to comply with the recommended maximum CPU core count guidelines published in the Exchange Team blog at: https://blogs.technet.microsoft.com/exchange/2015/10/15/ask-the-perf-guy-sizing-exchange-2016-deployments/.

All servers were configured identically, except that a processor comparison was designed and delivered. The only difference between the two configurations was that the servers had different processors, as identified below:

- HPE ProLiant DL380 Gen10 server (Gold 6132 processors, 2.6GHz, 14 cores) – run with 12 cores active per socket (24/28 total)
- HPE ProLiant DL380 Gen10 server (Gold 6126 processors, 2.6GHz, 12 cores)

Exchange sizing is often done using the SPECint2006 benchmark as an estimate of relative processor performance. As new processors reach the market, the new SPECint2006 Rate Value benchmark score is used to estimate how well it will perform in support of Exchange users. Microsoft provides a lookup or query tool, for this purpose: https://blogs.technet.microsoft.com/exchange/2012/04/30/released-processor-query-tool-v1-1/.

For example, inputting "Gold 6132" in Step 2 of the query tool yields an average result in Step 5 of 1614 when 28 cores is selected in the Step 4 drop-down. Note that Step 2 is sensitive to proper blank spacing before the 6132. The "SPEC score" of 1614 must be adjusted since the Exchange Server will be run with 12 cores active per socket (24/28 total) to remain compliant with the Exchange Preferred Architecture recommendation of 12 cores maximum per socket, (2 cores on each processor socket are disabled for the Gold 6132 14 core processors in the HPE ProLiant system configuration). The ratio of 24/28 cores for the 1614 score equals 1383 and this value is input in the Microsoft Exchange Server Role Requirements Calculator. The processor model "Gold 6126" looked up in the query tool or on the spec.org website yields a score of 1388 for 24 cores.

Sizing for failover scenario

It is a best practice to test the "worst-case" failover scenario, where the maximum amount of database copies are activated on the server. In this design, we tested a single server failure in the primary site, with no additional failover to the secondary data center mailbox servers.

According to Microsoft, the Preferred Architecture for Microsoft Exchange Server follows these design principles:

- Includes both high availability within the data center, and site resilience between data centers.
- Supports multiple copies of each database, thereby allowing for quick activation with clients shifted to other database copies.
- Reduces the cost of the messaging infrastructure.
- Increases Exchange server system availability by optimizing around failure domains and reducing complexity.

This Reference Architecture shows the configuration of the server, storage and Exchange databases meet the above solution requirements.

New security features in HPE ProLiant DL380 Gen10

Silicon root of trust and firmware protection

This new feature allows the firmware to be scanned and monitored through a series of integrity checks that initiate from an immutable link embedded in silicon. Furthermore, Hewlett Packard Enterprise has engineered the Gen10 servers with the ability to recover to a known good state in the unlikely event that firmware becomes compromised in some way.

Integrated Lights Out (iLO5) server management controller

This provides the silicon root of trust, ability to scan and monitor the chain of trust, and provide secure recovery. The iLO Advanced Premium Security Edition license that offers the highest level of commercial encryption capabilities, continual runtime detection of firmware validity, secure erase of the iLO5 NAND/NOR memory,, and secure recovery to authenticated states

HPE server options and management solutions

HPE Smart Array Controllers with Secure Encryption licenses help protect data at rest on attached storage devices. Unlike vendors that provide only self-encrypting drives, which often limits the number of protected drives. HPE offers controller-based encryption, so that all attached SAS and SATA drives are encrypted. This is a more cost-effective and comprehensive encryption solution for data at rest.

Enterprise Secure Key Manager (ESKM), a key management solution that works with Secure Encryption to provide centralized control and audit records for encryption keys. ESKM is FIPS 140-2 Level 2 validated and Common Criteria certified.

Secure components inside the server chassis

Trusted Platform Module (TPM) securely stores information needed to authenticate the server platform and to enable a measured boot process for the OS, which monitors the OS initialization process to see if the OS start-up has been compromised. TPM also supports specific capabilities such as Microsoft Windows BitLocker Drive Encryption. HPE offers both TPM 1.2 and 2.0 to support various operating systems.

Chassis Intrusion Detection Switch that detects if the chassis hood has been opened or closed and can send an alert through the iLO5 management device.

Spectre and Meltdown

In January 2018, two security vulnerabilities, Meltdown and Spectre, were unveiled by security researchers at Google's Project Zero in conjunction with academic and industry researchers from several countries. They are referred to as side-channel attacks, as they take advantage of the ability to extract information from instructions that have executed on a CPU using the CPU cache as side-channel. These are hardware bugs and are organized into 3 variants:

- Variant 1 (CVE-2017-5753, Spectre): Bounds check bypass
- Variant 2 (CVE-2017-5715, also Spectre): Branch target injection
- Variant 3 (CVE-2017-5754, Meltdown): Rogue data cache load, memory access permission check performed after kernel memory read

For guidance to mitigate speculative execution side-channel vulnerabilities see the following customer advisories:

- **Microsoft Security Advisory:** ADV180002 | Guidance to mitigate speculative execution side-channel vulnerabilities, https://portal.msrc.microsoft.com/en-us/security-guidance/advisory/ADV180002
- **Hewett Packard Enterprise Security Alert:** Side Channel Analysis Method allows information disclosure in Microprocessors, http://h22208.www2.hpe.com/eginfolib/securityalerts/SCAM/Side_Channel_Analysis_Method.html

Solution overview

The HPE ProLiant DL380 server family provides configurations that offer optimal combinations for Exchange Server and large, low-cost mailbox storage, with high-performance, reliability, and ease of deployment and management. The solution presented in this Reference Architecture is based on a building block using HPE ProLiant DL380 Gen10 servers each configured with 21 drives; 19 LFF HDDs, in a RAID-less JBOD configuration and 2 SFF SSDs, for the operating system and Exchange Transport files. The HPE ProLiant DL380 Gen10 server is available in a small form factor configuration as well.

See https://support.hpe.com/hpsc/doc/public/display?docId=emr_na-a00019684en_us&docLocale=en_US for information on the HPE ProLiant DL380 Gen10 server.

Look for Exchange Reference Architectures in the Hewlett Packard Enterprise Information Library at, http://h17007.www1.hpe.com/us/en/enterprise/reference-architecture/info-library/index.aspx?app=ms_exchange.

Using 10TB (7,200 RPM) SAS drives, this building block of 4 servers per site can provide 10,512 users with up to 21.5GB mailboxes with multiple database copies in a RAID-less JBOD configuration. This solution uses direct-attached storage (DAS) for the multiple database copies in the Database Availability Group (DAG) feature of Exchange Server 2016 to provide maximum database and service-level high availability. In addition to the resilient servers in the primary site, servers in a secondary site provide additional protection. All servers are designed with the same configuration for ease and consistency of deployment.

This Reference Architecture testing is designed to validate that the storage, CPU and memory subsystems can support this workload in both normal and peak operations, even in a failover scenario, where one of the servers in the primary site is offline. The solution is designed to withstand an outage of either a server within the site or the failover of the entire site and maintain availability to users. Proper sizing is done for the failover scenario, such that the CPU load is appropriate when the secondary site is unavailable and one of the servers in the primary site is offline, for either planned or unplanned downtime. During normal operations the CPU load will be substantially lower due to the headroom built into this failover model.

The goal is to have a working solution (and quorum) even with the loss of the entire secondary site (4 of the 8 servers). To maintain quorum we need to have more than half of voters on-line and able to vote. This DAG can use an external File Share Witness (FSW) server to act as a tiebreaker, which can be located in the primary site or in a third site. If you lose both the secondary site and an additional server in the primary site, the quorum is lost and the DAG goes offline without manual intervention. In order to return service, you need to manually activate the remaining servers by overriding DAG quorum. For more information on datacenter switchover in Exchange 2016 see: https://technet.microsoft.com/en-us/library/dd351049(v=exchg.160).aspx.

Identical configurations can be used for both the primary and secondary data center servers. This Reference Architecture compares the testing of two types of processors. Two sets of HPE ProLiant DL380 Gen10 servers were configured with 128GB of DDR4 RAM and either two (2) Intel Xeon Gold 6132 processors at 2.6GHz with 14 cores each, or two (2) Intel Xeon Gold 6126 processors at 2.6GHz with 12 cores each .

Note

To remain compliant with the Exchange Preferred Architecture recommendation of 12 cores maximum per socket, 2 cores on each processor (socket) are disabled for the Gold 6132 processors in the HPE ProLiant system configuration, details are shown later.

The HPE ProLiant DL380 Gen10 server with large form factor (LFF) drives model is shown in Figure 1 below. (The figure shows 300GB ENT drives instead of the 10TB MDL drives used in the testing).

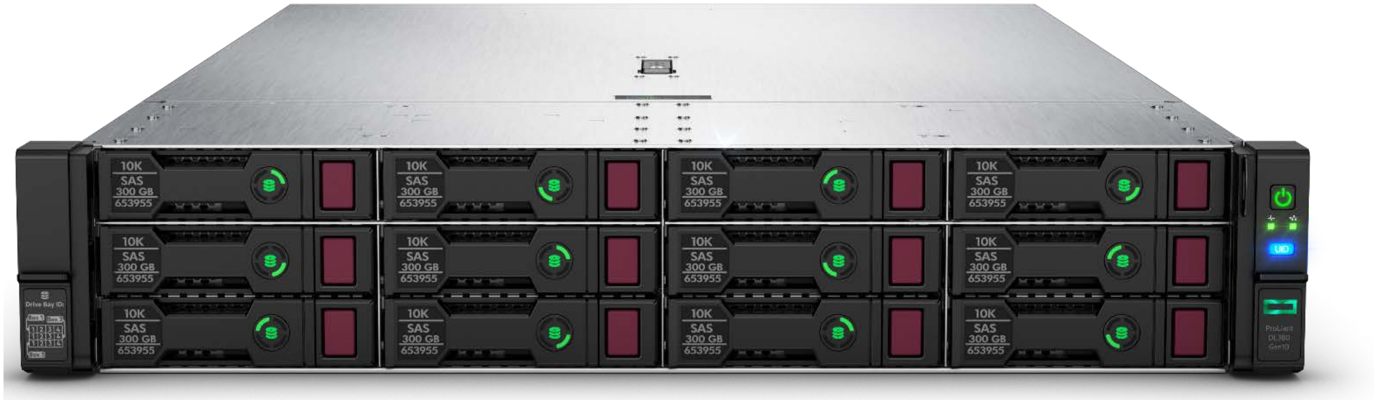

Figure 1. The HPE ProLiant DL380 Gen10 LFF server (Front view)

The HPE ProLiant DL380 Gen10 server uses an HPE Smart Array P816i-a controller with 4GB flash-backed write cache for internal storage. The disk storage for Exchange databases are HPE SAS large form factor drives of 10TB each.

- 2 (two) drives configured as RAID 1 for the operating system and Exchange Transport files (with a Windows partition for each)
- 18 (eighteen) 10TB large form factor (LFF) drives each configured as a single spindle RAID 0 for the Exchange databases
- 1 (one) 10TB large form factor (LFF) drive for either an Exchange recovery volume or AutoReseed volume, as recommended by Microsoft when using single drive RAID-less JBOD arrays

The bill of materials later in this document, Appendix A: Bill of materials, shows the configuration.

Note

For testing we used a RAID 1 pair of 800GB 12G Mixed Use Solid State Drives (SSDs) drives for the operating system and Exchange Transport partitions. Customers can easily substitute larger drives as recommended by the Microsoft Exchange Server Role Requirements Calculator or HPE Sizer for Exchange Server.

This storage configuration has been tested in accordance with the Microsoft Exchange Solution Reviewed Program (ESRP) presented at: https://technet.microsoft.com/en-us/office/dn756396.aspx. The solution is designed and tested to support 10,512 mailboxes at up to 21.5GB per mailbox for the primary mailboxes or personal archives when using 10TB drives. Mailbox space can be allocated between personal archive and primary mailbox. The difference between personal archives and primary mailboxes, and managing the content by policy is explained at: http://technet.microsoft.com/en-us/library/dd979795.aspx.

Each user is simulated at 150 messages per day (combined sends and receives), which represents an estimated 0.1005 IOPS per mailbox. This solution was tested at 0.1206 IOPS to allow for an additional 20% safety margin, as specified in the ESRP guidelines. The ESRP document published from the tested configuration is available at https://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=a00042439enw.

Figure 2 below provide details on the Exchange building block selected to support the development of the best practices outlined in this Reference Architecture.

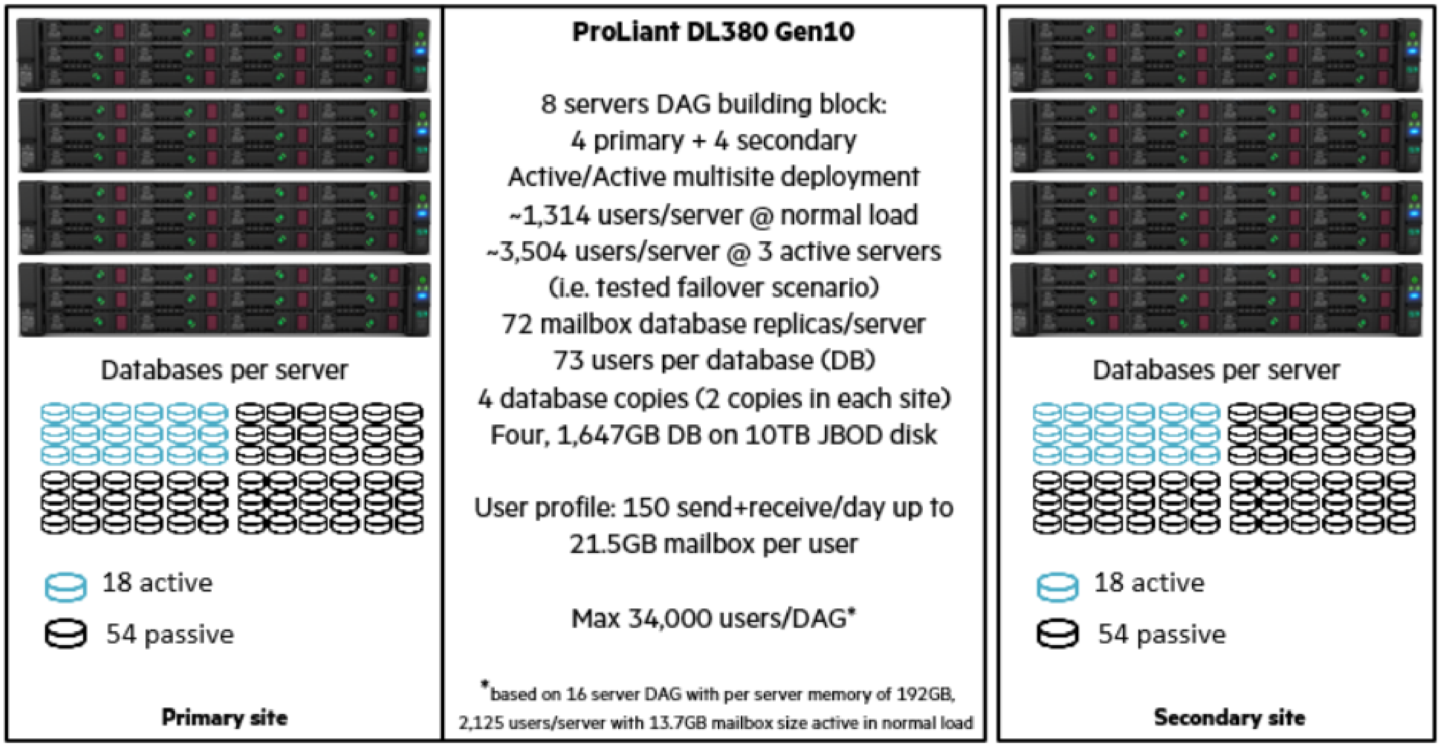

Figure 2 Exchange building block for primary and secondary data center

Understanding building blocks

This Exchange Server design uses a building block approach to adding servers to the Database Availability Group, thus the more servers that are added, the more the load can be spread over the additional servers. The total number of active users per server can increase as more servers are added. This Reference Architecture starts with a building block of 4 servers in each site, with each building block capable of serving the entire 10,512 users under the "worst-case" failover scenario. With 8 servers in each site, and up to 2,125 active users per server, the DAG is capable of supporting more users. However, it can only withstand the failure of one server before servers in the secondary (or DR) site must be used as the failover target. The choice ranges from the 4-server building block per site, up to 8 servers in each site with database copies evenly distributed, as the Microsoft Exchange Server Role Requirements Calculator demonstrates.

By increasing from 8 servers to a fully populated 16 servers-DAG, it is be possible to increase the number of users per server from 1,314 (10,512 per building block) to 2,125 per server, thus totaling 34,000 users in the DAG. The maximum mailbox size needs to be reduced, however, from 21.5GB to 13.7GB per user, due to the maximum of 100 active databases and a recommended maximum database size of 2TB.

Design principles

Exchange Server 2016 is deployed in a single role configuration, known as the Mailbox server. This means that the roles of Client Access Server (CAS) and Mailbox that were installed separately in previous versions of Exchange are combined on every Exchange 2016 server (with the exception of the Edge role, which is not involved here). Implementing Exchange servers as part of a DAG requires a network load balancer for high availability. This load balancing solution should also be highly available in order to maintain high availability of the overall Exchange service. The following link discusses load balancer solutions http://technet.microsoft.com/en-us/library/jj898588%28v=exchg.150%29.aspx. For more information refer to Microsoft Exchange Team Blog for load balancing in Exchange 2016 at https://blogs.technet.microsoft.com/exchange/2015/10/08/load-balancing-in-exchange-2016/.

Planning for quorum

Although beyond the scope of this paper, it is important that the placement of quorum or witness servers be understood. More information can be found under "Database availability group quorum models" at https://technet.microsoft.com/en-us/library/mt697595%28v=exchg.160%29.aspx.

Planning the secondary site

Configuration of the secondary data center or disaster recovery location and managing name resolution is a complex topic, very specific to the needs of your organization, and beyond the scope of this document. Please follow the guidance "High availability and site resilience" at https://technet.microsoft.com/en-us/library/mt697594(v=exchg.160).aspx.

HPE Sizer for Exchange Server 2016

The HPE Sizer for Exchange Server 2016, shown below in Figure 3, is a downloadable application allowing the end-user to customize input and receive an exact specification for servers and storage for Exchange. The configuration presented here was designed using the HPE Sizer for Exchange Server 2016. Additionally, the HPE Sizer for Exchange Server 2016 can be used to explore custom solutions or to make changes, using this Reference Architecture as a starting point. The test data presented here is also used by HPE to verify the accuracy of the results from the HPE Sizer for Exchange 2016.

The HPE Sizer for Exchange Server 2016 takes into account many factors such as email client usage profiles and mailbox size, including personal archives. The Sizer generates bills of material for various Exchange configurations, allowing the end-user to customize a solution design for their Exchange deployment. The Reference Architecture presented here can be customized using the HPE Sizer for Exchange Server 2016, which can be found at http://h20195.www2.hpe.com/v2/GetDocument.aspx?docname=4AA6-3720ENW.

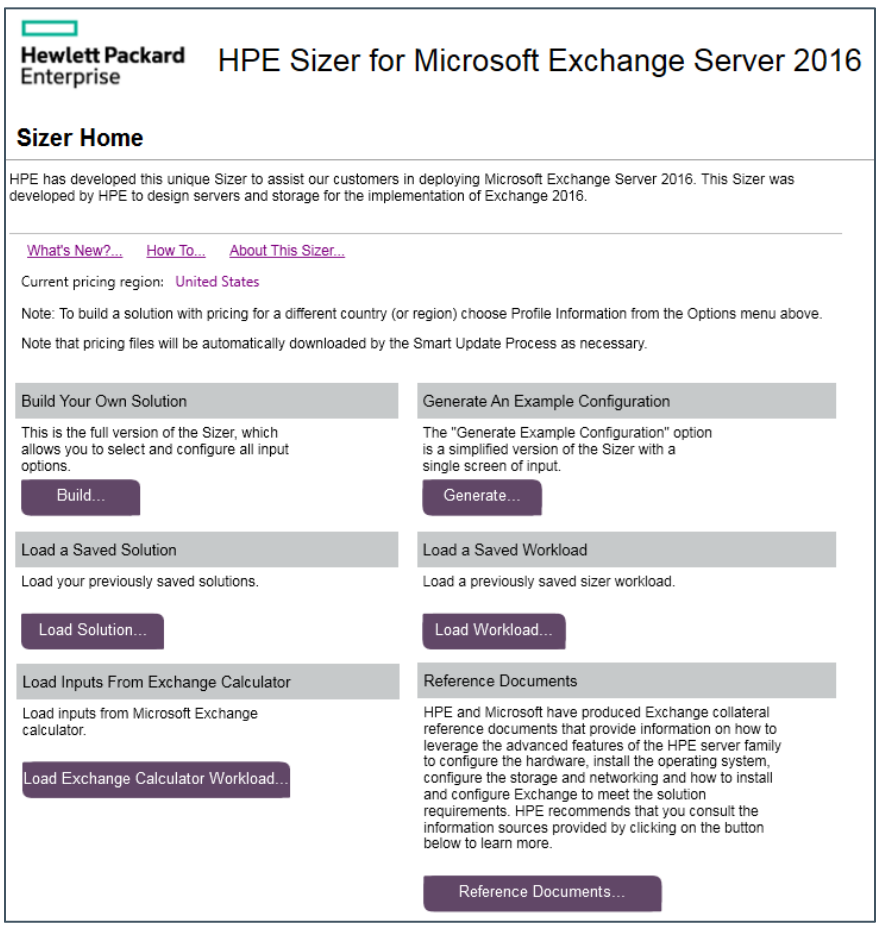

Figure 3. HPE Sizer for Exchange Server 2016

In addition to HPE Sizer for Exchange Server 2016, it is recommended that customers use the Microsoft Exchange Server Role Requirements Calculator. If this Calculator is used as a starting point, it can be imported into the HPE Sizer for Exchange Server 2016 using the button shown above in Figure 3 (Load Exchange Calculator Workload).

Figure 4 below shows the HPE Sizer output for Exchange Server 2016 on the HPE ProLiant DL380 Gen10 server, detailing the hardware configuration listed in Appendix A: Bill of material.

Note

The primary site hardware configuration is the same as the secondary site, which is not shown, to avoid repetition.

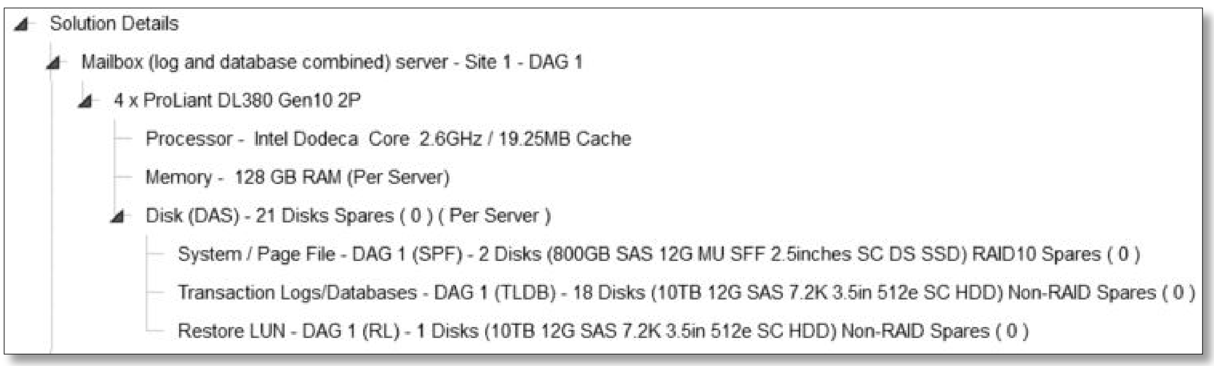

Figure 4. HPE Size output for Exchange Server 2016 on HPE ProLiant DL380 Gen10 server – primary site (secondary site is the same as the primary site)

The HPE Sizer for Exchange Server 2016 enables generating bills of material for various Exchange configurations, allowing the end-user to customize for their Exchange deployment. It takes into account many factors such as email client usage and mailbox size. The Sizer can also import the Microsoft Exchange Server Role Requirements Calculator (spreadsheet) as input, making it even easier to use the HPE Sizer for Exchange Server 2016.

Figure 5 is the database layout and failover from the Microsoft Exchange Server Role Requirements Calculator. The diagram below shows Secondary site is offline and a server (Srv04) in the Primary site is offline, with the databases active on the remaining servers.

			Primary site					Secondary site			
		Assigned	72	72	72	72		72	72	72	72
		Active	48	48	48	0		0	0	0	0
	Database Name	Active Server	Srv01	Srv02	Srv03	Srv04		Srv05	Srv06	Srv07	Srv08
Vol 1	DAG1-DB001	Srv01	1	2				3	4		
	DAG1-DB002	Srv02		1	2				3	4	
	DAG1-DB003	Srv03			1	2				3	4
	DAG1-DB004	Srv01	2			1		4			3
	DAG1-DB005	Srv01	3		4			1		2	
	DAG1-DB006	Srv02		3		4			1		2
	DAG1-DB007	Srv03	4		3			2		1	
	DAG1-DB008	Srv02		4		3			2		1
Vol 2	DAG1-DB009	Srv01	1			2		3			4
	DAG1-DB010	Srv02	2	1				4	3		
	DAG1-DB011	Srv03		2	1				4	3	
	DAG1-DB012	Srv03			2	1				4	3
	DAG1-DB013	Srv01	3	4				1	2		
	DAG1-DB014	Srv02		3	4				1	2	
	DAG1-DB015	Srv03			3	4				1	2
	DAG1-DB016	Srv01	4			3		2			1
Vol 3	DAG1-DB017	Srv01	1		2			3		4	
	DAG1-DB018	Srv02		1		2			3		4
	DAG1-DB019	Srv03	2		1			4		3	
	DAG1-DB020	Srv02		2		1			4		3
	DAG1-DB021	Srv01	3			4		1			2
	DAG1-DB022	Srv02	4	3				2	1		
	DAG1-DB023	Srv03		4	3				2	1	
	DAG1-DB024	Srv03			4	3				2	1
Vol 4	DAG1-DB025	Srv01	1	2				3	4		
	DAG1-DB026	Srv02		1	2				3	4	
	DAG1-DB027	Srv03			1	2				3	4
	DAG1-DB028	Srv01	2			1		4			3
	DAG1-DB029	Srv01	3		4			1		2	
	DAG1-DB030	Srv02		3		4			1		2
	DAG1-DB031	Srv03	4		3			2		1	
	DAG1-DB032	Srv02		4		3			2		1

Vol 5	DAG1-DB033	Srv01	1			2		3			4
	DAG1-DB034	Srv02	2	1				4	3		
	DAG1-DB035	Srv03		2	1				4	3	
	DAG1-DB036	Srv03			2	1				4	3
	DAG1-DB037	Srv01	3	4				1	2		
	DAG1-DB038	Srv02		3	4				1	2	
	DAG1-DB039	Srv03			3	4				1	2
	DAG1-DB040	Srv01	4			3		2			1
Vol 6	DAG1-DB041	Srv01	1		2			3		4	
	DAG1-DB042	Srv02		1		2			3		4
	DAG1-DB043	Srv03	2		1			4		3	
	DAG1-DB044	Srv02		2		1			4		3
	DAG1-DB045	Srv01	3			4		1			2
	DAG1-DB046	Srv02	4	3				2	1		
	DAG1-DB047	Srv03		4	3				2	1	
	DAG1-DB048	Srv03			4	3				2	1
Vol 7	DAG1-DB049	Srv01	1	2				3	4		
	DAG1-DB050	Srv02		1	2				3	4	
	DAG1-DB051	Srv03			1	2				3	4
	DAG1-DB052	Srv01	2			1		4			3
	DAG1-DB053	Srv01	3		4			1		2	
	DAG1-DB054	Srv02		3		4			1		2
	DAG1-DB055	Srv03	4		3			2		1	
	DAG1-DB056	Srv02		4		3			2		1
Vol 8	DAG1-DB057	Srv01	1			2		3			4
	DAG1-DB058	Srv02	2	1				4	3		
	DAG1-DB059	Srv03		2	1				4	3	
	DAG1-DB060	Srv03			2	1				4	3
	DAG1-DB061	Srv01	3	4				1	2		
	DAG1-DB062	Srv02		3	4				1	2	
	DAG1-DB063	Srv03			3	4				1	2
	DAG1-DB064	Srv01	4			3		2			1
Vol 9	DAG1-DB065	Srv01	1		2			3		4	
	DAG1-DB066	Srv02		1		2			3		4
	DAG1-DB067	Srv03	2		1			4		3	
	DAG1-DB068	Srv02		2		1			4		3
	DAG1-DB069	Srv01	3			4		1			2
	DAG1-DB070	Srv02	4	3				2	1		
	DAG1-DB071	Srv03		4	3				2	1	
	DAG1-DB072	Srv03			4	3				2	1

Vol 10	DAG1-DB073	Srv01	1	2			3	4		
	DAG1-DB074	Srv02		1	2			3	4	
	DAG1-DB075	Srv03			1	2			3	4
	DAG1-DB076	Srv01	2			1	4			3
	DAG1-DB077	Srv01	3		4		1		2	
	DAG1-DB078	Srv02		3		4		1		2
	DAG1-DB079	Srv03	4		3		2		1	
	DAG1-DB080	Srv02		4		3		2		1
Vol 11	DAG1-DB081	Srv01	1			2	3			4
	DAG1-DB082	Srv02	2	1			4	3		
	DAG1-DB083	Srv03		2	1			4	3	
	DAG1-DB084	Srv03			2	1			4	3
	DAG1-DB085	Srv01	3	4			1	2		
	DAG1-DB086	Srv02		3	4			1	2	
	DAG1-DB087	Srv03			3	4			1	2
	DAG1-DB088	Srv01	4			3	2			1
Vol 12	DAG1-DB089	Srv01	1		2		3		4	
	DAG1-DB090	Srv02		1		2		3		4
	DAG1-DB091	Srv03	2		1		4		3	
	DAG1-DB092	Srv02		2		1		4		3
	DAG1-DB093	Srv01	3			4	1			2
	DAG1-DB094	Srv02	4	3			2	1		
	DAG1-DB095	Srv03		4	3			2	1	
	DAG1-DB096	Srv03			4	3			2	1
Vol 13	DAG1-DB097	Srv01	1	2			3	4		
	DAG1-DB098	Srv02		1	2			3	4	
	DAG1-DB099	Srv03			1	2			3	4
	DAG1-DB100	Srv01	2			1	4			3
	DAG1-DB101	Srv01	3		4		1		2	
	DAG1-DB102	Srv02		3		4		1		2
	DAG1-DB103	Srv03	4		3		2		1	
	DAG1-DB104	Srv02		4		3		2		1
Vol 14	DAG1-DB105	Srv01	1			2	3			4
	DAG1-DB106	Srv02	2	1			4	3		
	DAG1-DB107	Srv03		2	1			4	3	
	DAG1-DB108	Srv03			2	1			4	3
	DAG1-DB109	Srv01	3	4			1	2		
	DAG1-DB110	Srv02		3	4			1	2	
	DAG1-DB111	Srv03			3	4			1	2
	DAG1-DB112	Srv01	4			3	2			1

Vol 15	DAG1-DB113	Srv01	1		2			3		4	
	DAG1-DB114	Srv02		1		2			3		4
	DAG1-DB115	Srv03	2		1			4		3	
	DAG1-DB116	Srv02		2		1			4		3
	DAG1-DB117	Srv01	3			4		1			2
	DAG1-DB118	Srv02	4	3				2	1		
	DAG1-DB119	Srv03		4	3				2	1	
	DAG1-DB120	Srv03			4	3				2	1
Vol 16	DAG1-DB121	Srv01	1	2				3	4		
	DAG1-DB122	Srv02		1	2				3	4	
	DAG1-DB123	Srv03			1	2				3	4
	DAG1-DB124	Srv01	2			1		4			3
	DAG1-DB125	Srv01	3		4			1		2	
	DAG1-DB126	Srv02		3		4			1		2
	DAG1-DB127	Srv03	4		3			2		1	
	DAG1-DB128	Srv02		4		3			2		1
Vol 17	DAG1-DB129	Srv01	1			2		3			4
	DAG1-DB130	Srv02	2	1				4	3		
	DAG1-DB131	Srv03		2	1				4	3	
	DAG1-DB132	Srv03			2	1				4	3
	DAG1-DB133	Srv01	3	4				1	2		
	DAG1-DB134	Srv02		3	4				1	2	
	DAG1-DB135	Srv03			3	4				1	2
	DAG1-DB136	Srv01	4			3		2			1
Vol 18	DAG1-DB137	Srv01	1		2			3		4	
	DAG1-DB138	Srv02		1		2			3		4
	DAG1-DB139	Srv03	2		1			4		3	
	DAG1-DB140	Srv02		2		1			4		3
	DAG1-DB141	Srv01	3			4		1			2
	DAG1-DB142	Srv02	4	3				2	1		
	DAG1-DB143	Srv03		4	3				2	1	
	DAG1-DB144	Srv03			4	3				2	1

Figure 5. Details of database layout and failover from the Microsoft Exchange Server Role Requirements Calculator

Solution components

HPE ProLiant DL380 solution summary

The following table provides an overview of the configuration for Microsoft Exchange Server 2016 and the tested HPE ProLiant DL380 Gen10 mailbox server building block.

Table 1. Tested HPE ProLiant DL380 Gen10 configuration details

Configuration	Detail
Number of users (mailboxes) per server • Normal operations • Any one site goes offline • during worst-case failover scenario (tested)	 • 1,314 • 2,628 • 3,504
Average user send/receive profile (and IOPS estimate)	150 messages per day (0.1005 base IOPS)
Maximum mailbox size	Up to 21.5GB (based on 10TB disks)
Database Availability Group (DAG) configuration • Total servers • Servers in primary / secondary locations • Copies in primary / secondary locations • Total users per DAG	 • 8 server building block (limit of 16 servers in a DAG) • 4 primary / 4 secondary • 2 copies in primary / 2 secondary • 10,512 users
HPE Mixed Use (MU) Hot-plug Solid State Drives (SSD)	• 2 drives (RAID 1) for Operating System (OS) and Transport volume are formatted with NTFS– relocate the Exchange Transport database queue and logs to a separate logical volume
Midline LFF drives	• 18 Database volumes are formatted with ReFS-- single drive RAID 0 arrays per server (10TB disk) • 1 Exchange recovery volume or AutoReseed volume – single drive RAID 0 array for emergency database operations or automatic replacement (10TB disk)

Best practices and configuration guidance

In addition to following the Microsoft Exchange 2016 Preferred Architecture and Exchange Product Team guidance, this section lists the additional configuration options to consider and test before deploying Exchange 2016 into your production environment.

Server power mode

The default power mode for the HPE ProLiant server is Dynamic Power Savings Mode. The recommendation from the Microsoft Exchange product team is to set the server BIOS to allow the operating system to manage power, and use the "High performance" power plan in Windows® (https://blogs.technet.microsoft.com/exchange/2015/04/30/troubleshooting-high-cpu-utilization-issues-in-exchange-2013). Because of the constant high workload of the Exchange mailbox role, the power mode was changed, as shown below in Figure 6, to OS Control Mode for testing and the server is configured for Maximum Performance.

Note

Changing the server power mode from the default of “Dynamic Power Savings Mode” to “OS Control Mode” has a significant benefit in performance and should not be overlooked unless the end user validates that the power saving profile can sustain the production workload.

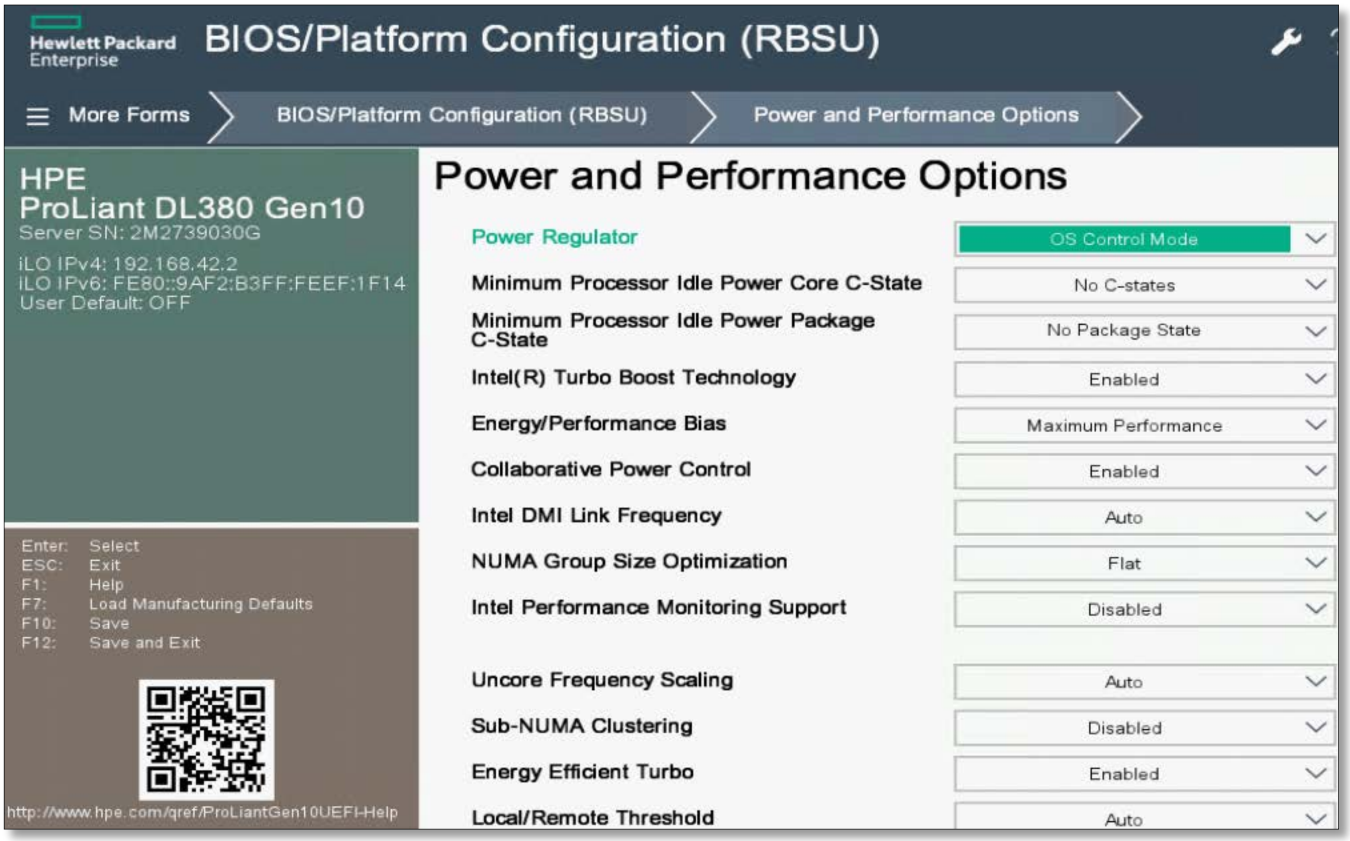

Figure 6. Changing the server power mode to OS Control Mode

Network subsystem configuration

In configuring the network subsystem, the network requirements from Microsoft should be adhered to. Refer Network requirements for more information: http://technet.microsoft.com/en-us/library/dd638104.aspx#NR.

The solution here follows the Microsoft Preferred Architecture simplified approach of using a single 10GbE network interface for both client traffic and server to server replication traffic. Since a single drive RAID 0 array (RAID-less JBOD design with many databases increases the chance for database re-seeding (when a single disk fails).

An additional network may be desirable for server management including remote lights-out management. As shown below in Figure 7 the HPE ProLiant DL380 Gen10 server is equipped with HPE FlexFabric 10GbE 2-port 534FLR-SFP+ Adapter to provide 10GbE network interface and a separate HPE Integrated Lights-Out 5 (iLO 5) to provide remote deployment and management features for the server nodes. Note that iLO 5 operates at 1GbE via the dedicated iLO 5 port.

Figure 7. Rear view of example HPE ProLiant DL380 Gen10 server highlighting iLO5 port

Deploying Exchange Server 2016

The detailed steps to prepare the environment and install Exchange are well documented on Microsoft TechNet and are not duplicated here. Since this was a “greenfield” deployment, and not a migration or interoperability design, the process may be significantly different than in other organizations. Each server is prepared with the Exchange Server 2016 prerequisites (https://technet.microsoft.com/en-us/library/bb691354(v=exchg.160).aspx and https://technet.microsoft.com/en-us/library/aa996719(v=exchg.160).aspx), and the Exchange 2016 Cumulative Update (CU) 8 was installed. See https://technet.microsoft.com/en-us/library/jj907309(v=exchg.160).aspx for most recent updates.

To achieve optimal performance within the RA, the following guidelines and best practices must be considered:

- **Hyper-Threading:** Also known as simultaneous multithreading (SMT), Hyper-Threading should be turned off when Exchange 2016 is run directly on physical servers according to Microsoft at http://blogs.technet.com/b/exchange/archive/2013/05/06/ask-the-perf-guy-sizing-exchange-2013-deployments.aspx. It is acceptable to enable Hyper-Threading on physical hardware that is hosting hypervisor software with Exchange virtual machines, so you may see it enabled in other Reference Architectures from Hewlett Packard Enterprise. But capacity planning should always be done based on physical CPUs and not total SMT or Hyper-Threaded logical processors (typically double the number of physical processors but certainly not doubling the performance).
- **Processor cores:** In order to remain compliant with the Exchange Preferred Architecture limit of 24 cores per server, the HPE ProLiant BIOS/Platform Configuration (RBSU) system option is used to disable the additional cores on the server with the Gold 6132 processors. This procedure was not necessary for the 12 core Gold 6126 processors.

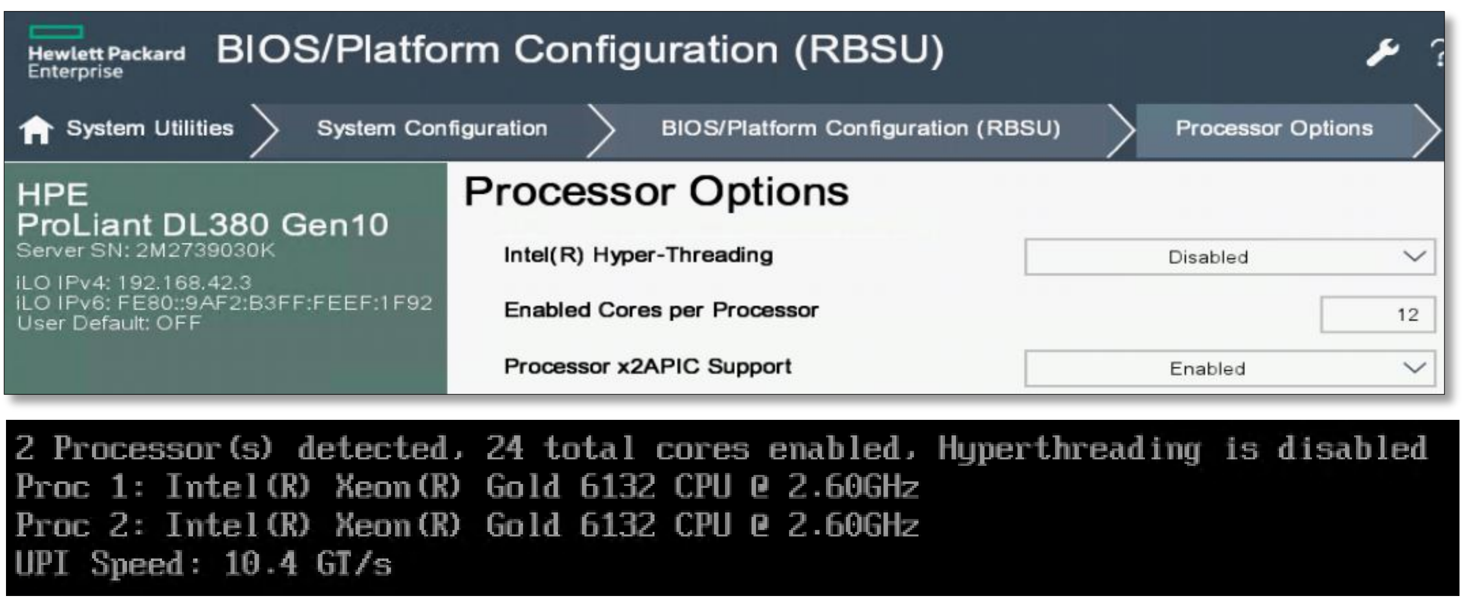

Figure 8. To remain compliant with the Exchange Preferred Architecture, cores on each processor are limited to 12 in the BIOS/Platform Configuration (RBSU)

- **Windows volumes:** Windows volumes used for operating system and transport volumes are formatted with NTFS, and Windows volumes used for Exchange mailbox databases should be formatted with ReFS and the 64KB allocation unit size set for best performance and data integrity features must be disabled. The AutoReseed component that automatically allocates and formats spare disks is called the Disk Reclaimer. Note that that if the disk has been used within 24 hours as the emergency recovery volume then AutoReseed won't use it (automatically via Disk Reclaimer). For more information on Disk Reclaimer see https://technet.microsoft.com/en-us/library/dn789209(v=exchg.160).aspx. For more information on ReFS and storage options see: https://technet.microsoft.com/en-us/library/ee832792%28v=exchg.150%29.aspx (Exchange 2013 version, with no update for 2016).

Note

ReFS is supported with Storage Spaces, Storage Spaces Direct, and non-removable direct attached drives. ReFS is not supported with hardware virtualized storage such as SANs or RAID controllers in non-passthrough mode. USB drives are also not supported.

Database paths need to be consistent across all servers in the DAG that have a copy of the mailbox database, thus volume mount points are used and each drive is mounted under C:\ExchangeDatabases\ or any path that is preferred. However, Exchange AutoReseed functionality and the Desired State Configuration use a specific structure. See: http://blogs.technet.com/b/mhendric/archive/2014/10/17/managing-exchange-2013-with-dsc-part-1-introducing-xexchange.aspx (Exchange Server 2013 version, no update at time of publication for Exchange Server 2016).

An example PowerShell script is shown below, as it was used in the Reference Architecture configuration. Be sure to validate it in your own test environment before attempting to use it, for example limiting the number of drives to a single one, instead of 1 through 18 as shown below.

```
1..18 | foreach {md "c:\ExchangeDatabases\$_" -erroraction ignore;
Initialize-Disk $_ -PartitionStyle GPT -PassThru -erroraction ignore;
New-Partition -DiskNumber $_  -UseMaximumSize -erroraction ignore |
Add-PartitionAccessPath -AccessPath "C:\ExchangeDatabases\$_" -PassThru |
Format-Volume -FileSystem REFS -SetIntegrityStreams:$false -Confirm:$false ;
Set-Partition -DiskNumber $_ -NoDefaultDriveLetter:$True -Partition 2}
```

- **Storage (Disk) encryption:** HPE Smart Array Controllers with Secure Encryption licenses protect data at rest on attached storage devices. Unlike vendors that provide only self-encrypting drives, which often limits the number of protected drives, HPE Smart Array encryption was used to encrypt each disk in this RA, providing data encryption. This was done before adding the single drive RAID 0 arrays, so that each new disk was automatically encrypted by policy. This is a more cost-effective and comprehensive encryption solution for data at rest.
- **Database Availability Group (DAG):** This solution is built upon the Database Availability Group (DAG) resiliency feature in Exchange 2016. This feature is the base component of HA and site resilience framework built into Exchange 2016. A DAG is a group of 2 to 16 mailbox servers that each host a set of database copies and provide database-level recovery from failures that affect individual servers or databases.
- Windows Server 2012 R2 introduced the ability to create a DAG that does not need an administrative access point, for example using the command such as below, which does not specify an IP address:

```
New-DatabaseAvailabilityGroup -Name DAG1 -DatabaseAvailabilityGroupIPAddresses
([System.Net.IPAddress])::None -WitnessServer [Server] -WitnessDirectory C:\DAG1
```

- **File Share Witness directory settings:** The Database Availability Group (DAG) leverages file share witness to maintain the quorum. The witness server should be the member of the domain and it cannot be a DAG member. The File Server role is installed on the witness server and the Exchange Trusted Subsystem group is added to the local Administrator group on the witness server. For more information on the witness directory settings see: https://technet.microsoft.com/en-us/library/dd297985(v=exchg.160).aspx.

Pre-staging the DAG computer account

As noted in the Microsoft Exchange guidance "Managing database availability groups" at https://technet.microsoft.com/en-us/library/dd298065(v=exchg.150).aspx, for environments where computer account creation rights are restricted, or where computer accounts are created in containers other than the default computers container, it is highly advisable to pre-stage and assign permissions on the Cluster Name Object (CNO). Create a computer account for the CNO and either assign full control to the computer account of the first mailbox server you are adding to the DAG, or assign full control to the Exchange Trusted Subsystem group. This ensures the necessary security context. It is necessary first to select View, Advanced Features to see the Security tab on the object, and then change the Object Types in the search to include Computers so that you can select the Exchange Server account when setting permissions. Failure to configure the CNO may result in an error that may be confusing, such as "the fully qualified domain name for (CNO) could not be found".

As shown in Figure 9, below, adding server nodes to the DAG cluster will install the necessary Windows Failover Clustering components.

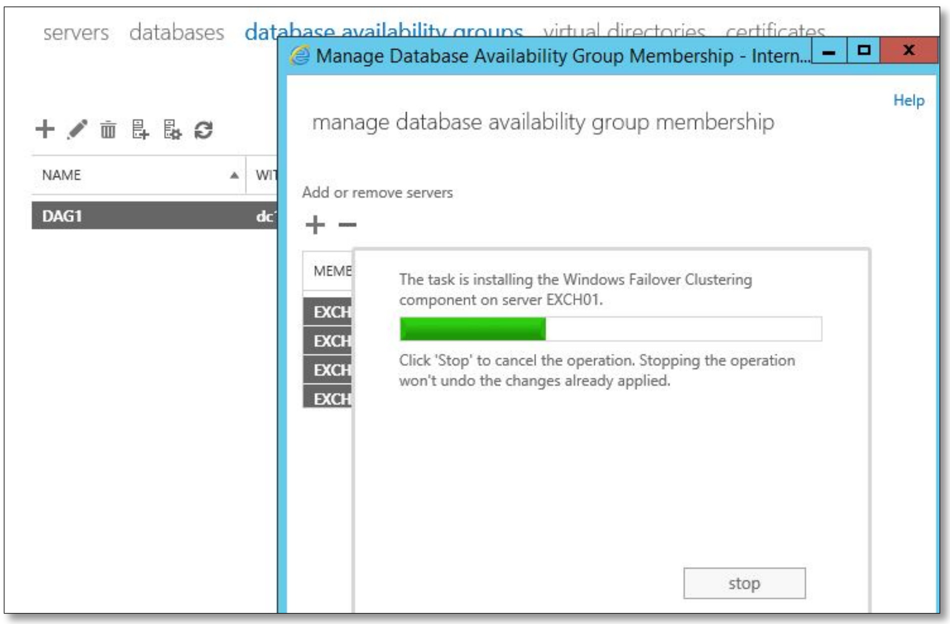

Figure 9. Adding server nodes to the DAG cluster

Transport location

A new volume is created to move the Exchange Transport files off of the C:\ operating system volume. This was done as part of the installation, using the script in "C:\Program Files\Microsoft\Exchange Server\V15\Scripts" and running the following in the Exchange Management Shell (not Windows PowerShell).

```
.\Move-TransportDatabase.ps1 -queueDatabasePath 'T:\Queue' -queueDatabaseLoggingPath 'T:\Queue\Logs'
```

If the Exchange transport load is expected to be typical, the transport storage volume can share the same physical drives as the OS/boot volume, but a separate RAID 1 array is desirable for maximum performance. The Microsoft Exchange Server Role Requirements Calculator can help make this determination, and also takes into consideration the capacity and performance constraints of the OS/boot volume.

Antivirus (antimalware) and anti-spam

Exchange 2016 offers built-in anti-spam and anti-malware protection, but customers may choose to deploy a third-party product installed on the Exchange servers. No difference was observed in previous comparison testing between the performance of the built-in antivirus protection and a third-party product. Product selection can be made on desired features instead, as the built-in antivirus protection is not a full-featured product. The built-in antivirus protection is disabled using the commands below, to make the change take effect.

```
& $env:ExchangeInstallPath\Scripts\Disable-Antimalwarescanning.ps1
```

```
Restart-Service MSExchangeTransport
```

According to the following article, “antispam agents are available in the Transport service on Exchange 2016 Mailbox servers, but they are not installed by default” -- see https://technet.microsoft.com/en-us/library/bb201691(v=exchg.160).aspx.

Exchange Server Performance Health Checker Script

The Microsoft Exchange Server Performance Health Checker Script checks common configuration settings such as product versions, pagefile settings, power plan settings, NIC settings, as well as processor and memory information. It works on Exchange 2010, 2013 and 2016. Example output for the Exchange servers is shown below in Figure 10. It is recommended to run the Exchange Server Health Checker Script periodically to ensure that your environment is operating at peak health and that configuration settings have not been inadvertently changed. Download the script at: https://gallery.technet.microsoft.com/office/Exchange-2013-Performance-23bcca58

```
Exchange Health Checker version 2.18
System Information Report for EXCH03 on 05/03/2018 23:17:18


Hardware/OS/Exchange Information:
        Hardware Type: Physical
        Manufacturer: HPE
        Model: ProLiant DL380 Gen10
        Operating System: Microsoft Windows Server 2016 Datacenter
        Exchange: Exchange 2016 CU8
        Build Number: 15.1.1415.2
        Server Role: Mailbox
        MAPI/HTTP Enabled: True
Pagefile Settings:
        Pagefile Size: 32778
.NET Framework:
        Version: Windows Server 2016 .NET 4.6.2
Power Settings:
        Power Plan: High performance
Http Proxy Setting:
        Setting: <None>
NIC settings per active adapter:
        Interface Description: HPE FlexFabric 10Gb 2-port 534FLR-SFP+ Adapter  #45 [Embedded FlexibleLOM 1 Port 1]
                Driver Date: 2017-09-08
                Driver Version: 7.13.105.0
                Link Speed: 10000 Mbps
                RSS: Enabled
Processor/Memory Information
        Processor Type: Intel(R) Xeon(R) Gold 6132 CPU @ 2.60GHz
        Hyper-Threading Enabled: No
        Number of Processors: 2
        Number of Physical Cores: 24
        Number of Logical Cores: 24
        NUMA Group Size Optimization: BIOS Set to Flat
        Megacycles Per Core: 2601
        Physical Memory: 128 GB

TCP/IP Settings:
The TCP KeepAliveTime value is configured optimally (900000)

LmCompatibilityLevel Settings:
        LmCompatibilityLevel is set to: 3
        LmCompatibilityLevel Description: Clients use only NTLMv2 authentication, and they use NTLMv2 session security if the s
erver supports it. Domain controllers accept LM, NTLM, and NTLMv2 authentication.
        LmCompatibilityLevel Ref: https://technet.microsoft.com/en-us/library/cc960646.aspx

Hotfix Check:
KB3206632 is Installed
```

Figure 10. Output from the Exchange Server Performance Health Checker Script

Storage controller configuration

The HPE Flexible Smart Array P816i-a/4GB FBWC was used where the controller, and cache settings were left at the default, 10% read and 90% write. Previous testing has shown that modifying this setting has little benefit, so the default was tested to see if it performed adequately. However, follow the caution below and do not use HBA mode, which has the effect of disabling all write cache on the controller. The effect should be obvious if someone has incorrectly built the server using HBA mode, as all disk writes will be slow (e.g., greater than 10ms).

A feature in HPE Smart Storage Administrator (SSA) allows configuring individual drives as RAID 0 arrays, which has the same net result of presenting the drives directly to the operating system, as shown below in Figure 11. Either way, using scripting or the HPE SSA User Interface (UI), the configuration is much simpler than in the past, and no Smart Array scripting is necessary to configure disks as single drive RAID 0 arrays (RAID-less JBOD). The only scripts necessary are to use diskpart or PowerShell to configure and format the Windows volumes. An example of the script or commands used in testing is shown earlier in this document.

The HPE Smart Array P816i-a storage controller also offers a choice of Power Mode. Testing was performed with the default Max Performance, since Exchange is a storage intensive application. Any deviation from this setting may degrade performance so test and apply any changes only with due caution.

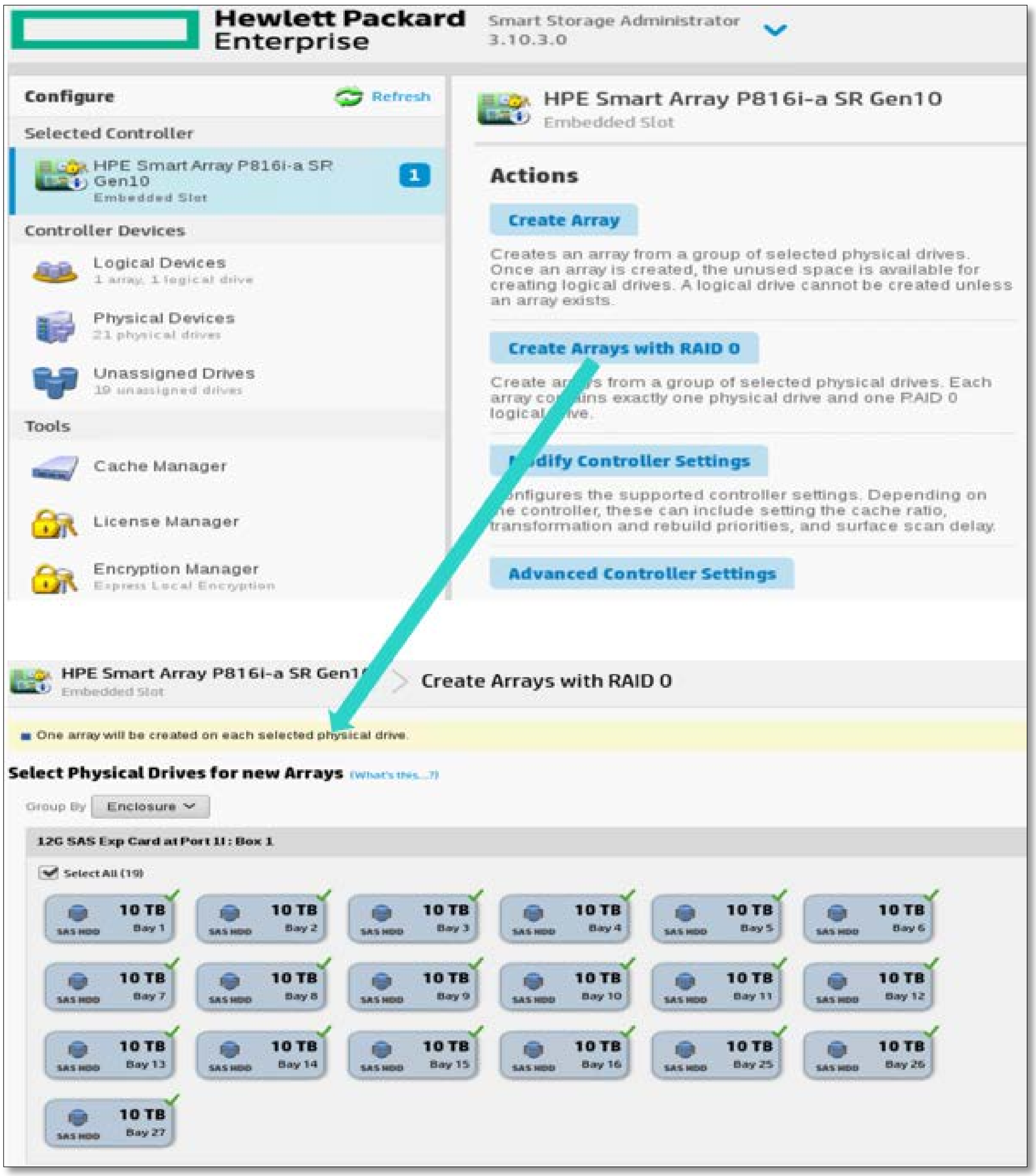

Figure 11. HPE Smart Array P816i-a controller configuration for single drive RAID 0 arrays (RAID-less JBOD)

Capacity and sizing

Exchange Server 2016 test methodology

When evaluating the scalability and performance of an Exchange building block in a lab environment, our engineers use tools provided by Microsoft to generate a simulated Exchange workload on the systems and analyze the effect of that workload. One tool, Jetstress, is designed for disk subsystem testing, and the other tool, LoadGen, evaluates server performance with a given set of client messaging and scheduling actions generating server load and mail flow within the organization.

Microsoft Exchange Server Jetstress tool

Microsoft Exchange Server Jetstress tool works with the Exchange Server 2016 database engine to simulate the Exchange database, log and replication load on disks. Jetstress helps verify the performance and stability of a disk subsystem before putting an Exchange 2016 server into production. Jetstress simulates only Exchange database load for a specific number of users, or determines the maximum load at given latency thresholds.

After successful completion of the Jetstress disk performance (2 hour) and stress (24 hour) tests, an Exchange Server 2016 disk subsystem is determined to be adequately sized (in terms of the performance criteria you establish) for the user count and profiles you selected. For more information on Jetstress, visit: microsoft.com/en-us/download/details.aspx?id=36849. Note that Jetstress 2013 works with Exchange 2016 CU4 by providing the required ESE files from the Exchange Server 2016 installation source, as documented in Jetstress.

Microsoft Exchange Server Load Generator tool

Microsoft Exchange Server Load Generator tool (LoadGen) produces a simulated client workload against a test Exchange deployment. This workload evaluates how Exchange performs and is used to analyze the effect of various configuration changes on Exchange behavior and performance while the system is under load. It is capable of testing both Exchange 2013 and 2016.

LoadGen simulates the delivery of multiple MAPI client messaging requests to an Exchange server. To simulate the delivery of these messaging requests, you run Exchange Load Generator tests on client computers. These tests send multiple message requests to the Exchange server, which causes a mail load.

After the tests are complete, use the results to perform the following tasks:

- Verify that the server and storage configuration performed are within design specifications.
- Identify bottlenecks on the server.
- Validate Exchange settings and server configuration.
- Verify that client load was accurately represented.

For more information on LoadGen, visit: microsoft.com/en-us/download/details.aspx?id=40726.

Analysis and recommendations

To understand how the solution performs, the Reference Architecture was tested using both Microsoft Exchange Server Jetstress and Microsoft Exchange Server Load Generator tools to demonstrate server and storage performance.

Jetstress was used to test the storage performance by simulating the Exchange 2016 database and transaction log workloads while measuring disk performance to determine achievable read and write IOPS. The results validated that the Reference Architecture was adequately sized to support 10,512 mailbox users with a 150 message-user work profile with additional performance headroom even in a failover scenario where only one Exchange server is online.

The second phase of testing used LoadGen to simulate client load and validate the server configuration. Test results were then analyzed to verify that each server was adequately sized to support the mailbox users with a 150 messages per user, per day user work profile with additional performance headroom.

Since each test tool provides a different function it is important to use both. The database volumes are sized for four 1.6TB databases on each hard drive, and Jetstress created four 1.8TB databases on each of the 18 volumes, thus filling the system towards its designed maximum capacity. LoadGen cannot easily create such large databases and even initializing for the average mailbox size (shown below in Figure 12) takes weeks for 10,512 users. However LoadGen does not need such large databases to provide stress testing and validation of the CPU, memory and network resource consumption. In addition, Jetstress databases cannot be used for network replication testing, but LoadGen creates actual

Exchange Server databases and copies can be created on other Exchange servers. These are all industry-accepted best practices for using these Exchange test tools.

Another area of focus for testing is the location of the Exchange Transport files, and standard LoadGen testing may not provide the stress seen in organizations with heavy inbound and outbound mail flow, in addition to expansion and delivery of many Distribution Lists (DLs).

```
Count    : 3504
Average  : 309.922
Sum      : 1085969
Maximum  : 594
Minimum  : 220
```

Figure 12. Tested average mailbox size (shown in megabytes) for a single server, calculated during LoadGen testing

Storage performance results

The storage performance testing, using Jetstress, exercises the storage with the targeted sustainable Exchange I/O load for 2 hours. The test measures storage I/O response times under the Exchange I/O load and has several thresholds for a passing score, such as average log writes under 10ms and average database read latencies under 20ms. The data below is the sum of all logical disk I/O and an average of all the logical disk I/O latency in the 2-hour test duration.

These storage performance results have been previously published in an ESRP document at: http://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=a00042439enw.

Scenario tested: Failure scenario with three servers active in primary site

Figure 13, below shows a failover scenario (planned or unplanned failover) running the active databases on 3 of the 4 servers in our primary site in an active/active deployment, with no over to the secondary site. Notice that 48 of 72 databases are active on each server. Our design uses 18 HDD in RAID-less JBOD, thus 12 drives have 3 copies active and the 6 drives have 2 copies, i.e. 12x3 + 6x2 = 48 databases. The Jetstress validation test (such as for the ESRP) results shows that the Achieved Transactional I/O per Second in the test report higher than the Total Database Required IOPS / Server predicted in the Microsoft Exchange Server Role Requirements Calculator and the I/O Database Reads Average Latency <20ms.

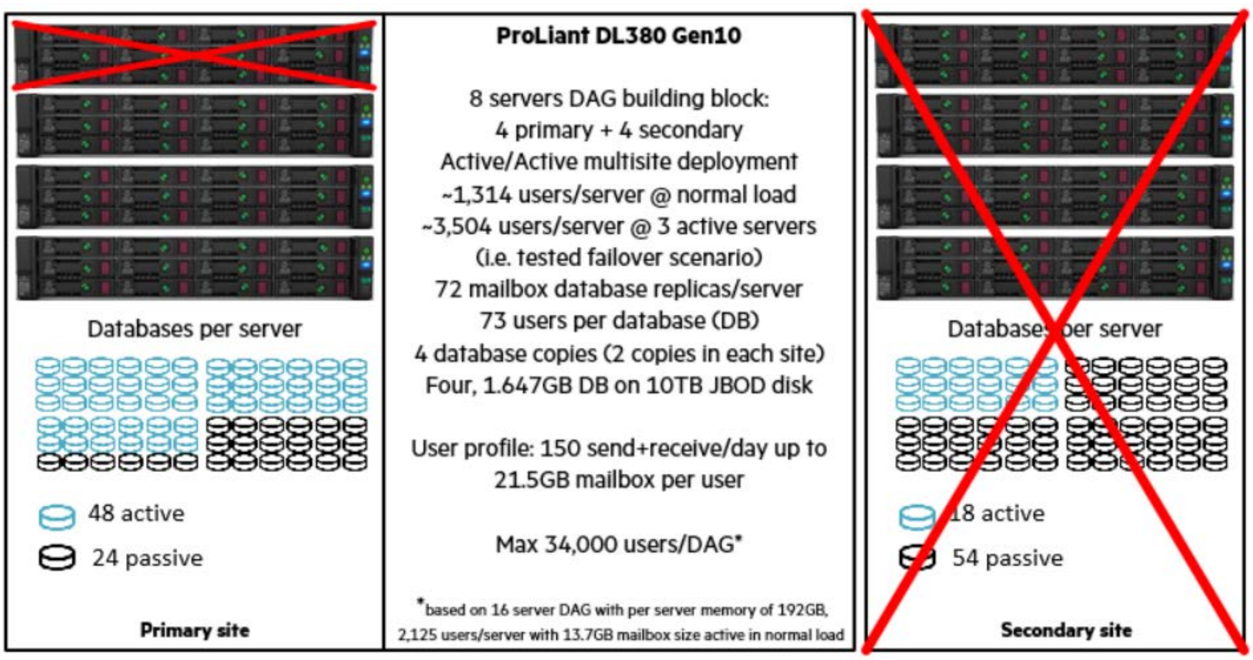

Figure 13. Simplified architectural diagram showing Exchange database copies in an Active/Active site design for Worst-case failure operation

Table 2. Storage performance results via Jetstress – single server summary

Single server summary	Performance results
Database I/O	
Needed Disk Transfers/sec*	634
Achieved Database Disk Transfers/sec	1695
Database Disk Reads/sec	1181
Database Disk Writes/sec	514
Average Database Disk Read Latency (ms)	18.11
Average Database Disk Write Latency (ms)	0.54
Transaction Log I/O	
Log Disk Writes/sec	118
Average Log Disk Write Latency (ms)	0.12

* This row represents the IOPS necessary to satisfy the IOPS per mailbox required for the number of active databases in this solution.

There are 2 database backup/recovery performance test reports in this part of the testing. The first one measures the sequential read rate of the database files and the second measures the recovery performance (replaying transaction logs into the database). One important characteristic of these tests is that the read rates and log replay rates are measured with all planned active databases for the designed failover scenario under concurrent or simultaneous load.

The database read-only performance test measures the maximum rate at which databases can back up via Volume Shadow Copy Service (VSS) aware backup applications. The following table shows the average and sum of database reads.

Table 3. Jetstress test results – Database backup

Database backup	Performance results
Average MB read/sec per database	41.04
MB read/sec total	2955

LoadGen performance results

All LoadGen testing presented here was performed with the server power set to OS Control Mode as advised earlier in this document. In this Reference Architecture Exchange Server 2016 performance testing is done for two configurations of HPE ProLiant DL380 Gen10 server. The only difference between the two servers was the use of different processors as identified below:

- HPE ProLiant DL380 Gen10 server (Gold 6132 processors, 2.6GHz, 14-cores) – run with 12 cores active per socket (24/28 total)
- HPE ProLiant DL380 Gen10 server (Gold 6126 processors, 2.6GHz, 12-cores)

The following table summarizes the test results of an 8- hour workload (150 send/receive, with Outlook 2007 Cached mode). The total of messages submitted and delivered per user, per day, reaches 150 over the measured interval, thus the server has ample processor and is running at 38% CPU (Gold 6132) or 41% CPU (Gold 6126), which indicates that it can sustain this load.

Table 4. Summary of LoadGen results for eight hour workload (150 send/receive, Outlook 2007 Cached simulation) per server

Eight hour workload test	Performance results (Gold 6132 processor with 24/28 cores)	Performance results (Gold 6126 processor total 24 cores)
Disk reads/sec	1130	1415
Disk writes/sec	388	405
MSExchange RPC Client Access		
User count	3438	3437
MSExchangeIS Store		
Messages Delivered /user/day	122	123
Messages Submitted/user/day	30	30
RPC Averaged Latency (ms)	0.58	0.77
RPC Operations/sec	3,228	3,287
Network Interface		
MBytes received/sec	3.22	2.76
MBytes sent/sec	3.74	2.46
Average CPU utilization	38%	41%

The load of 150 messages/day represents the *average* load per mailbox, but peak times can frequently exceed that average, so a generally accepted practice is to test at twice the average level to simulate peak usage. Tests were run at an effective rate of 300 messages/day to simulate this peak impact and are shown below in Table 6. The total of messages submitted and delivered per user, per day, reaches 300 over the measured peak interval, thus the server has ample processor and is only running at 56% CPU (Gold 6132) or 56% CPU (Gold 6126), which indicates that it can sustain this peak load.

Table 5. Summary of LoadGen results for four hour peak workload (Outlook 2007 Cached simulation) per server

Four hour peak workload test	Performance results (Gold 6132 processor with 24/28 cores)	Performance results (Gold 6126 processor total 24 cores)
Disk reads/sec	1131	1730
Disk writes/sec	639	651
MSExchange RPC Client Access		
User count	3439	3439
MSExchangeIS Store		
Messages Delivered/user/day	246	244
Messages Submitted/user/day	61	61
RPC Averaged Latency (ms)	0.64	0.92
RPC Operations/sec	6,232	6,241
Network Interface		
MBytes received/sec	5.83	4.85
MBytes sent/sec	6.38	4.79
Average CPU utilization	56%	56%

Again, the performance is quite satisfactory, including the ability to handle the peak performance (2x IOPS), in this test scenario with all servers and databases online. This indicates there is headroom for one planned or unplanned failover during normal operations.

Summary

This white paper demonstrated a validated solution design for Exchange Server 2016 on the HPE ProLiant DL380 Gen10 servers. The workload demonstrates that the HPE ProLiant DL380 Gen10 solution adequately handles the normal and peak loads for 3,504 users per server with three of four servers in one site of the building block active. The CPU utilization is sufficiently below the calculation by the Microsoft Exchange Server Role Requirements Calculator, which indicates that the server design can sustain usage profile and peak usage patterns. The HPE Sizer for Exchange Server 2016 can help assess your environment, allowing modification of the user profile to match your needs.

The HPE ProLiant DL380 Gen10 server as configured here contains 19 large form factor (LFF) hard disk drives 10TB each. By using what is known as single drive RAID-less JBOD (single drive RAID 0 arrays), we use the entire capacity of the 10TB drives for several Exchange databases (up to 4, given a maximum set size of 1.8TB databases). The design and testing shown here in this document covers the importance of understanding how each of up to 4 database copies on a single disk are activated in a failover scenario. It is very important to note that only 2 or 3 copies will be activated on any given disk in the failover scenario where 3 of 4 servers are active and all databases are online.

This solution is designed using a building block approach, with a multi-copy database design using Exchange Database Availability Groups (DAGs). This Reference Architecture testing is designed to validate that the storage, CPU and memory subsystems can support this workload in both normal and peak operations, even in a failover scenario, where one of the servers in the primary site is offline. The solution is designed to withstand an outage of either a server within the site or the failover of the entire site and maintain availability to users. Proper sizing is done for the failover scenario, such that the CPU load is appropriate when the secondary site is unavailable and one of the servers in the primary site is offline, for either planned or unplanned downtime. During normal operations the CPU load will be substantially lower due to the headroom built into this failover model.

This design is based on a specific number of servers per DAG which determines the number of users that can be hosted on a single server in a failover scenario. For example, 8 servers (4 servers per site) are sized with 1,314 users per server in normal operations, and 16 servers in the DAG can run up to 2,125 users per server in normal operations (with smaller mailboxes due to the increased user count). The number of servers per DAG is a flexible design decision that customers can make using the tools, resources and design considerations described in this document.

All servers were configured identically, except that a processor comparison was designed and delivered. The only difference between the two configurations was the servers had different processors, as identified below:

- HPE ProLiant DL380 Gen10 server (Gold 6132 processors, 2.6GHz, 14 cores) – run with 12 cores active per socket (24/28 total)
- HPE ProLiant DL380 Gen10 server (Gold 6126 processors, 2.6GHz, 12 cores)

Exchange sizing is often done using the SPECint2006 benchmark as an estimate of relative processor performance. As new processors reach the market, the new SPECint2006 Rate Value benchmark score is used to estimate how well it will perform in support of Exchange users.

Detailing the test configuration and workload proof points can help simplify and expedite the deployment of this solution, or similar solutions, based on a building block approach. Deploying an enterprise-ready Exchange Server 2016 solution can be an involved and time consuming process. This white paper detailed the hardware necessary to support 10,512 users with 21.5GB mailboxes with a messaging profile of 150 emails sent or received, per user, per day, in a highly available and disaster resilient solution on Hewlett Packard Enterprise servers with internal large form factor drives and HPE Smart Array P816i-a controller. This white paper also provided highlights of installing and configuring the server and storage, and configuring Exchange Server 2016 to meet the solution requirements.

Finally, performance results from both the Jetstress and LoadGen test tools show this solution meets the requirements in the CPU, RAM, and storage subsystems in a high availability failover scenario where only one server is online to handle the load. The test data presented here is also used by Hewlett Packard Enterprise to verify the accuracy of the results from the HPE Sizer for Exchange Server 2016.

Spectre and Meltdown performance impact for Microsoft Exchange server 2016 deployment

The Spectre and Meltdown system ROM update was installed on the HPE ProLiant DL380 Gen10 Servers. The mailbox servers provided ample processor headroom for scaling peak workloads during LoadGen solution validation testing.

Implementing a proof-of-concept

As a matter of best practice for all deployments, Hewlett Packard Enterprise recommends implementing a proof-of-concept test environment that matches as closely as possible the planned production environment. In this way, appropriate performance and scalability characterizations can be obtained. For help with a proof-of-concept, contact an HPE Services representative (hpe.com/us/en/services/consulting.html) or your Hewlett Packard Enterprise partner.

Appendix A: Bill of materials

Note

Part numbers are at time of publication and subject to change. The bill of material does not include complete support options or other rack and power requirements. If you have questions regarding ordering, please consult with your HPE Reseller or HPE Sales Representative for more details. hpe.com/us/en/services/consulting.html

Table 6. Bill of Materials

Qty	Part number	Description
		Rack
2	BW908A	HPE 642 1200mm Shock Intelligent Rack
		Configuration 1 – Servers using Gold 6126 Processor
4	868705-B21	HPE DL380 Gen10 12LFF CTO Server
4	826862-L21	HPE DL380 Gen10 6126 Xeon-G FIO Kit
4	826862-B21	HPE DL380 Gen10 6126 Xeon-G Kit
16	815100-B21	HPE 32GB 2Rx4 PC4-2666V-R Smart Kit
4	826685-B21	HPE DL380 Gen10 3LFF Rear SAS/SATA Kit
4	826686-B21	HPE DL38X Gen10 4LFF MID-plane HDD
8	872376-B21	HPE 800GB SAS 12G MU SFF SC DS SSD
76	857644-B21	HPE 10TB SAS 7.2K LFF SC He 512e DS HDD
4	700751-B21	HPE FlexFabric 10Gb 2-port 534FLR-SFP+ Adapter
4	804338-B21	HPE Smart Array P816i-a SR Gen10 Ctrlr
4	870549-B21	HPE DL38X Gen10 12Gb SAS Expander
4	875241-B21	HPE 96W Smart Storage Battery 145mm Cbl
4	864279-B21	HPE TPM 2.0 Gen10 Kit (Optional)
8	865438-B21	HPE 800W FS Ti Ht Plg LH Pwr Sply Kit
4	Q2F26AAE	HPE Smart Array Secure Encryption E-LTU
4	JG081C	HPE X240 10G SFP+ SFP+ 5m DAC Cable
		Configuration 2 – Servers using Gold 6132 Processor
4	868705-B21	HPE DL380 Gen10 12LFF CTO Server
4	826870-L21	HPE DL380 Gen10 6132 Xeon-G FIO Kit
4	826870-B21	HPE DL380 Gen10 6132 Xeon-G Kit
16	815100-B21	HPE 32GB 2Rx4 PC4-2666V-R Smart Kit
4	826685-B21	HPE DL380 Gen10 3LFF Rear SAS/SATA Kit
4	826686-B21	HPE DL38X Gen10 4LFF MID-plane HDD
8	872376-B21	HPE 800GB SAS 12G MU SFF SC DS SSD
76	857644-B21	HPE 10TB SAS 7.2K LFF SC He 512e DS HDD

4	700751-B21	HPE FlexFabric 10Gb 2-port 534FLR-SFP+ Adapter
4	804338-B21	HPE Smart Array P816i-a SR Gen10 Ctrlr
4	870549-B21	HPE DL38X Gen10 12Gb SAS Expander
4	875241-B21	HPE 96W Smart Storage Battery 145mm Cbl
4	864279-B21	HPE TPM 2.0 Gen10 Kit (Optional)
8	865438-B21	HPE 800W FS Ti Ht Plg LH Pwr Sply Kit
4	Q2F26AAE	HPE Smart Array Secure Encryption E-LTU
4	JG081C	HPE X240 10G SFP+ SFP+ 5m DAC Cable

Resources and additional links

HPE Solutions for Exchange, http://h17007.www1.hpe.com/us/en/enterprise/reference-architecture/info-library/index.aspx?app=ms_exchange

HPE Sizer for Microsoft Exchange Server 2016, http://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=4AA6-3720ENW

HPE Reference Architectures, hpe.com/info/ra

HPE Servers, hpe.com/servers

HPE Storage, hpe.com/storage

HPE Advisory and Transformation Services, hpe.com/us/en/services/consulting.html

To help us improve our documents, please provide feedback at hpe.com/contact/feedback.

Learn more at hpe.com/info/ProLiant Servers

Sign up for updates

Hewlett Packard Enterprise

a00048480enw, June 2018