Improving Microsoft SQL Server Database performance with HPE Persistent Memory on HPE ProLiant DL380 Gen9

Technical white paper

Contents

Executive summary

The demands of database implementations continue to escalate. Faster transaction processing speeds, scalable capacity, and increased flexibility are required to meet the needs of today's business. At the same time, enterprises are looking for cost-effective, open-architecture, industry standard solutions that don't include vendor lock-in or carry the high price tag attached to proprietary solutions.

Recently introduced HPE Persistent Memory brings ground-breaking storage performance that helps businesses meet those demands.

HPE Persistent Memory products deliver the performance of memory with the persistence of traditional storage. Customers are looking for offerings that enable faster business decisions and the HPE Persistent Memory modules deliver outstanding performance to put data to work more quickly in your business. It is ideal for accelerating database and analytics workloads

Microsoft® developed Windows® operating system drivers and SQL Server changes to make use of HPE Persistent Memory modules (NVDIMMs). This Reference Configuration will demonstrate the ease of configuring HPE 8GB NVDIMM modules along with the performance gains achievable in a Microsoft SQL Server 2016 Database environment.

HPE NVDIMMs can be leveraged several ways within a SQL Server 2016 environment:

- File storage on storage class memory (SCM) using traditional NTFS block devices
- File acceleration using Storage Spaces and storage class memory as a Direct Access (DAX) device cache to SSD
- Log acceleration using SQL Server 2016 Persistent Memory Manager (PMM)

Testing showed a 35% transactional throughput increase by just implementing PMM logging.

Together these improvements stack to provide better performance and lower processor utilization resulting in either increase workload co-location or licensing savings (by deploying the workload on fewer cores). Our tested configuration would require 63% less core licenses than compared to the highest core processor.

This Reference Configuration focuses on both the file acceleration and log acceleration use cases of HPE Persistent Memory and SQL Server 2016.

Target audience: This Hewlett Packard Enterprise white paper is designed for IT professionals who use, program, manage, or administer large databases that require high availability and high performance. Specifically, this information is intended for those who evaluate, recommend, or design new IT high performance architectures.

This white paper describes testing completed in June 2016.

Document purpose: The purpose of this document is to describe a Reference Configuration, highlighting benefits and key implementation details to technical audiences.

Solution overview

HPE Persistent Memory

To gain a real competitive advantage, you need to enable faster business decisions. The HPE Persistent Memory portfolio delivers outstanding performance to put data to work more quickly in your business. HPE Persistent Memory offerings are not just new hardware technology, but a complete software ecosystem designed to work with today's applications and workloads, including databases and analytics workloads.

The HPE 8GB NVDIMM module is the first offering in the HPE Persistent Memory product category. It delivers the performance of memory with the resiliency you have come to expect from HPE storage technology. Customers can have confidence that business-critical data is safe because HPE utilizes higher endurance DRAM and components that help verify data is moved to non-volatile technology in the event of a power loss.

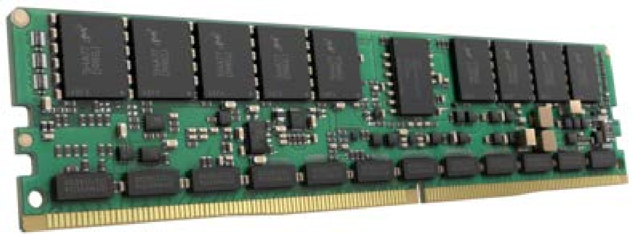

Figure 1. HPE 8GB NVDIMM module

HPE NVDIMMs achieve performance improvements by providing low latency I/O and by reducing I/O path overhead due to direct access. The following figure illustrates the simplified paths achieved in block and direct access modes. The reduction in overhead compared to traditional block access directly results in reduced kernel time, freeing CPU cycles for additional user mode activities (more transactions) or potentially reducing platform core capacity needed for the same workload.

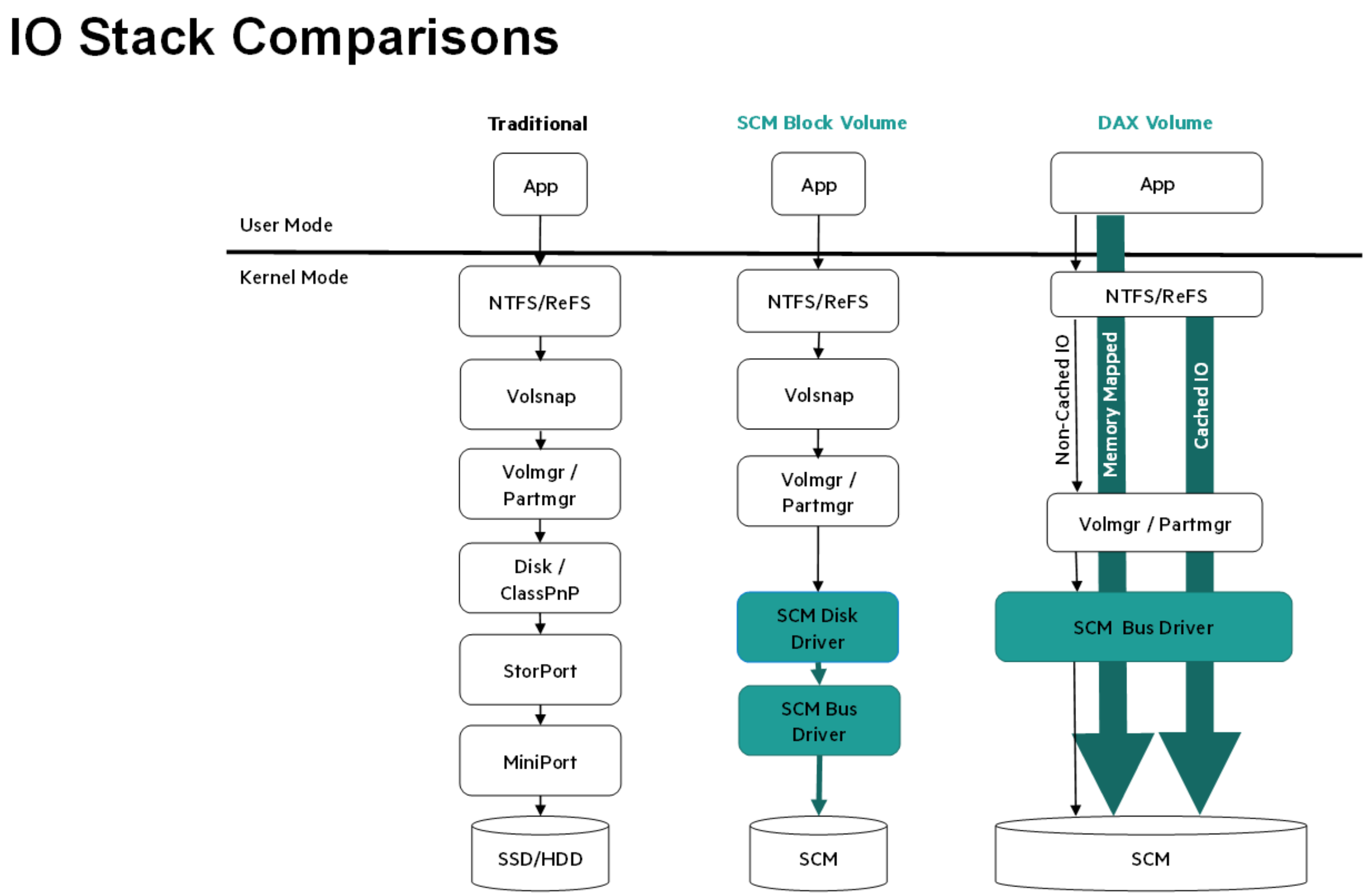

Figure 2. I/O stack comparisons

HPE Persistent Memory offers the following features:

- Turbo-charged performance delivering up to 4x faster transaction performance

  HPE ProLiant DL360 Gen9 and DL380 Gen9 servers equipped with HPE 8GB NVDIMM modules increase performance for write intensive workloads delivering up to 2x+ faster database logging performance[1]. The NVDIMM modules are designed to speed customer application workloads, delivering up to 4x+ faster OLTP replication functions[2] enabling faster workloads.

- Technology designed to make your business data resilient

  The HPE Persistent Memory modules include a flash component plus an HPE Smart Storage Battery that provides you with a persistent storage capability at memory speeds without the data volatility of memory.[3] Active data runs on the DRAM component of NVDIMM which not only provides outstanding performance but also offers greater endurance than traditional storage media.[4]

- Solutions designed around your business workloads

  HPE Persistent Memory is designed around industry applications and workloads to deliver the performance of memory with the persistence of storage. A complete hardware and software ecosystem provides a comprehensive persistent memory solution for your business.

Solution components

This white paper provides configuration and performance information for HPE 8GB NVDIMMs in a SQL Server 2016 RTM Database environment. The tests were run on an HPE ProLiant DL380 Gen9 server running Windows Server® 2016 Technical Preview 5.

The HPE ProLiant DL380 Gen9 server was configured with the following components:

- Two frequency-optimized 8-core Intel® Xeon® E5-2667 v4 processors at 3.2 GHz
- 256GB memory (8 x 32GB RDIMMs)
- 16 x HPE 8GB NVDIMM modules
- 8 x 400GB 12G SAS write-intensive SSD drives
- 2 HPE SX350 1600GB PCIe Workload Accelerators (Data)
- 1 HPE SX350 1300GB PCIe Workload Accelerator (Logs)

[1] Internal HPE lab testing on an HPE ProLiant DL380 Gen9 E5 2600 v4 with HPE 8GB NVDIMM-N, Dec 2015
[2] Internal HPE lab testing on an HPE ProLiant DL380 Gen9 E5 2600 v4 with HPE 8GB NVDIMM-N, Dec 2015
[3] Based on the NVDIMM utilizing NAND Flash as a persistent store and the HPE Smart Storage Battery providing the backup power source to move data from DRAM to NAND Flash
[4] Endurance comparison based on comparing Program Erase Cycles of DRAM to Program Erase Cycles of NAND Flash. DRAM can have up to 10 trillion more program erase cycles than NAND Flash.

Figure 3 shows the location of the NVDIMMs (outlined in red) and RDIMMS, along with the key disk configuration of the HPE ProLiant DL380 Gen9 server.

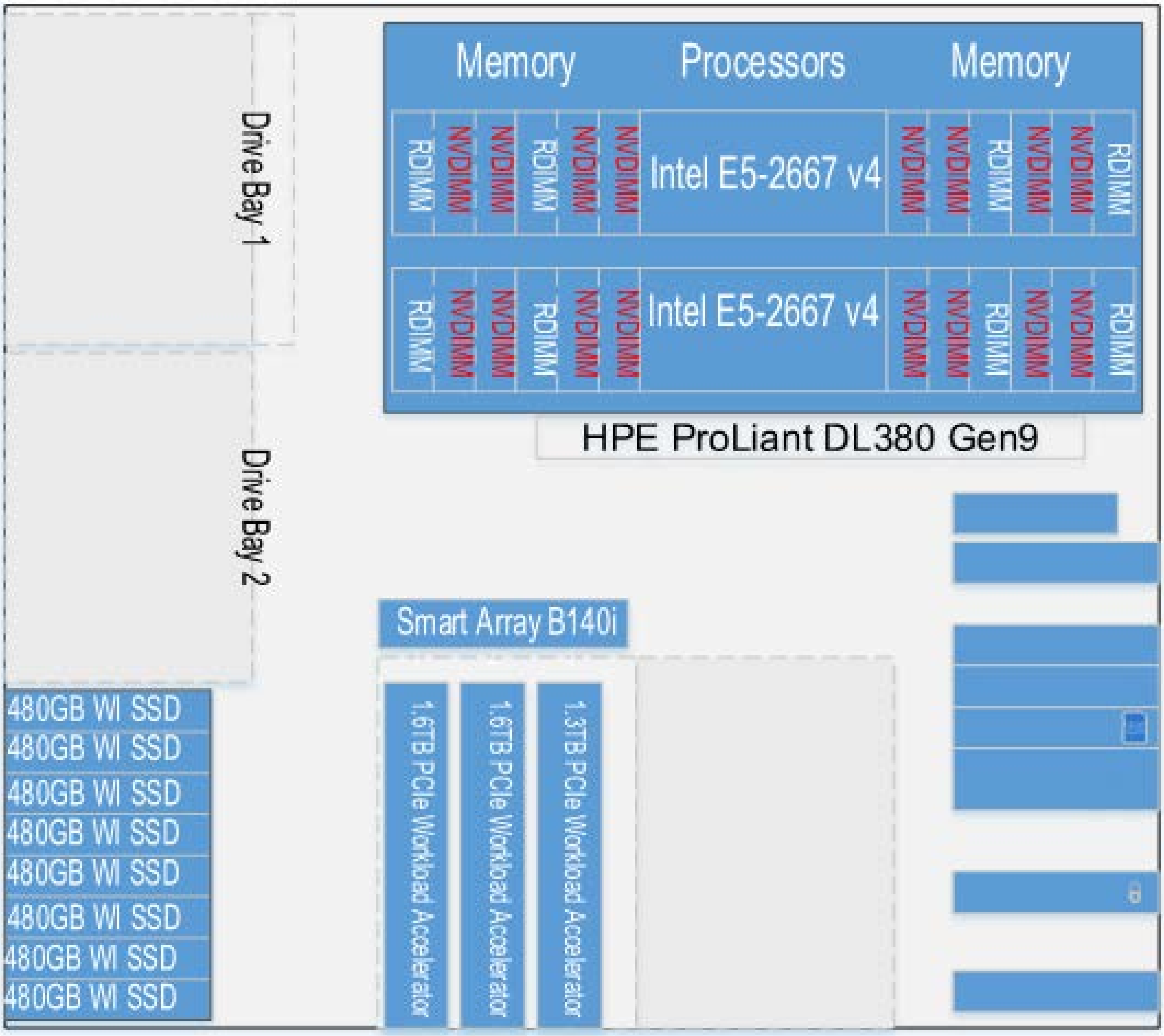

Figure 3. Diagram of memory and disk layout for HPE ProLiant DL380 Gen9 server

Software

- Windows Server 2016 Technical Preview 5
- SQL Server 2016 RTM

Best practices and configuration guidance for the SQL Server Databases using HPE NVDIMMs

HPE ProLiant DL380 Gen9 BIOS

- Hyper-Threading—Enabled
- Intel Turbo Boost—Enabled
- HPE Power Profile—Maximum Performance

Physical NVDIMM configuration best practices

- The NVDIMMs must be configured according to the population rules outlined in the QuickSpecs and HPE NVDIMM User Guide. See Appendix B for the configuration used for this testing.
- Balance the total memory capacity across all processors.
- Only RDIMMs can be mixed with HPE 8GB NVDIMMs. No other memory types may be used when Type 1 NVDIMMs are present.
- Interleave the NVDIMM modules (via BIOS settings) to create one block device on each socket. See the HPE NVDIMM User Guide for interleaving instructions.

Storage configuration for SQL Server database logs using SQL Server PMM (Tail of the Log enhancement) and HPE NVDIMM

1. Prepare and verify NVDIMM is ready for SQL Server PMM use:
   a. Bring NVDIMM disk online using disk management tool
   b. Format drive for NTFS

      Execute: `Format <drive>: /y /q /DAX`
   c. Verify volume is a DAX volume: `fsutil fsinfo volumeinfo <drive>:`

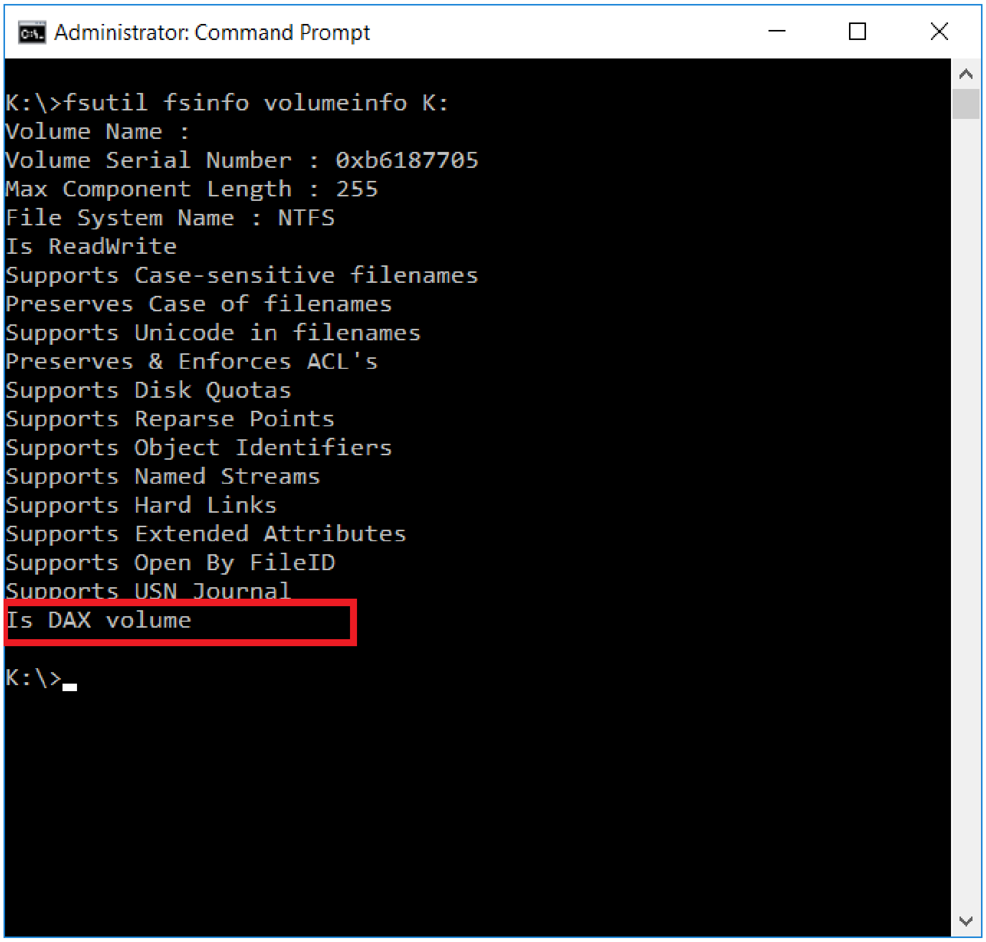

Figure 4. Fsutil command to verify DAX status of NVDIMM

2. Enable trace flag -T9921
3. Add a special secondary log file for each database (tail of the log):

```
USE [master]
GO

ALTER DATABASE [tpcc_05] ADD LOG FILE ( NAME = N'TOL_05', FILENAME = N'K:\tail_05.ldf' , SIZE = 179200KB , FILEGROWTH = 0)
ALTER DATABASE [tpcc_06] ADD LOG FILE ( NAME = N'TOL_06', FILENAME = N'K:\tail_06.ldf' , SIZE = 179200KB , FILEGROWTH = 0)
ALTER DATABASE [tpcc_07] ADD LOG FILE ( NAME = N'TOL_07', FILENAME = N'K:\tail_07.ldf' , SIZE = 179200KB , FILEGROWTH = 0)
ALTER DATABASE [tpcc_08] ADD LOG FILE ( NAME = N'TOL_08', FILENAME = N'K:\tail_08.ldf' , SIZE = 179200KB , FILEGROWTH = 0)
GO
```

Storage configuration for SQL Server database data using Storage Spaces Directs (S2D) and HPE NVDIMM

HPE NVDIMMs are used in Storage Spaces to provide a write-back cache for SSD media. In addition to the cache performance advantage, the disks can be configured with "mirror" resiliency for added reliability. In our test scenario two disks are mirrored and the NVDIMMS are mirrored as part of the pool configuration.

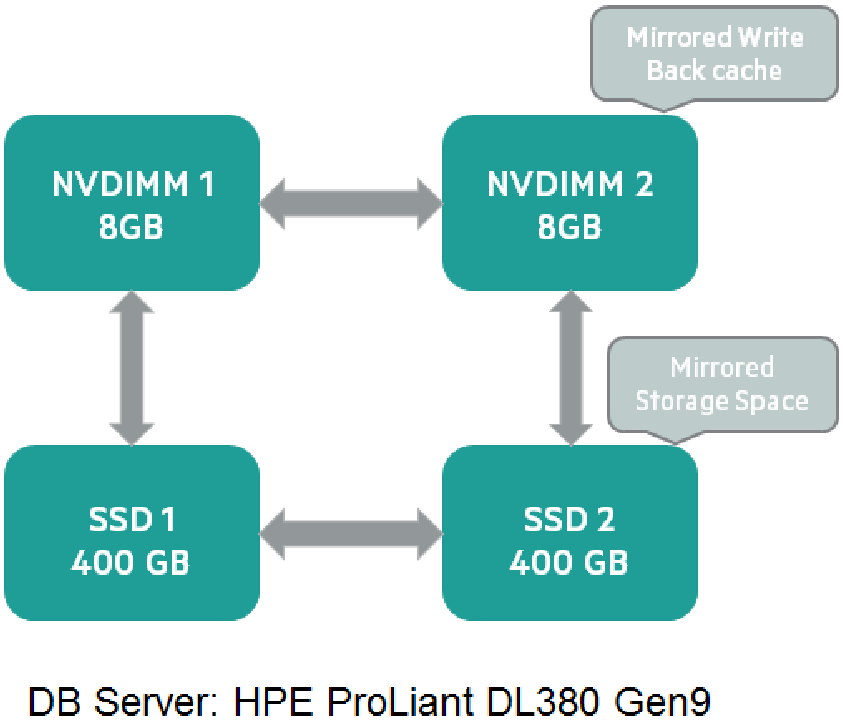

Figure 5. Typical Storage Spaces disk layout

PowerShell configuration steps for implementing Storage Spaces pool with HPE NVDIMM:

1. Identify NVDIMMs and SSDs:

```
$pd = get-physicalDisk -canpool $true
```

2. Get two physical SSD disks, set media type to HDD and clear them:

```
$disk = get-disk -number [disk id]
set-physicaldisk -uniqueid $disk.uniqueid -mediatype HDD
clear-disk -number [disk id]-RemoveOEM -RemoveData
```

3. Create a new storage pool with the two disks and two NVDIMMs:

```
new-storagepool -StorageSubSystemFriendlyName Storage -FriendlyName SCM_Pool -PhysicalDisks $pd
```

4. Set NVDIMM usage to "Journal":

```
Get-PhysicalDisk -FriendlyName 2c* | Set-PhysicalDisk -Usage Journal
```

5. Create the Storage Spaces disk:

```
New-VirtualDisk -StoragePoolFriendlyName SCM_Pool_Disk -FriendlyName SCM_Mirror -
ResiliencySettingName Mirror -UseMaximumSize -ProvisioningType Fixed
```

6. Initialize disk:

```
Get-VirtualDisk -FriendlyName SCM_Pool_Disk |Initialize-Disk-PartitionStyle GPT"
```

7. Format the filesystem:

```
$partition = New-Partition -DiskNumber 6 -AssignDriveLetter -UseMaximumSize
Format-Volume -Partition $partition -FileSystem NTFS
```

SQL Server 2016 configuration best practices

SQL Server settings used in this reference configuration:

- Ensure MAXDOP for instance and databases are set to 1
- “Lock pages in memory” right granted to user running instance, to prevent buffer pool memory release
- Trace flag -T834 added for large page allocation
- Trace flag -T9921 added to enable SQL Server persistent memory logging (PMM).
- Maximum SQL Server memory was set well below physical max to prevent OS paging to disk.

Capacity and sizing

Workload description

The Microsoft SQL Server 2016 workload was tested using HammerDB, an open-source tool. The tool implements an OLTP-type workload (60 percent read and 40 percent write) with 8k random reads and writes (during checkpoints) and 64k sequential writes to the log file.

Four 240GB OLTP databases were used to achieve a balanced 4:1 database size to memory ratio, resulting in a highly CPU and moderately I/O intensive mixed workload. The workload was run until buffer pool was warm (64k data read ahead drops to 8k) and transactional throughput was measured with the Microsoft SQL Server Perfmon counter “Batch requests per second”.

SQL Server was deployed on both write-intensive SSD and HPE PCIe Workload Accelerator media types to obtain NVDIMM transactional improvement measurements. This was measured using the Microsoft SQL Server “Batch requests per second” Perfmon counter among others.

The workload was un-paced in that each user executed transactions back to back as fast as the system responded. 15 users were created per database to bring the overall system utilization to the 75-80% range prior to enabling Microsoft PMM.

Workload results

Two different storage configurations were tested using Microsoft SQL Server 2016 Persistent Memory Manager (PMM):

- Four databases deployed on write-intensive SSD drives
- Four databases deployed on HPE Workload Accelerators

Write-Intensive SSD drives

Figure 6 shows the transactional gain when PMM logging is enabled on databases deployed on write-intensive SSD drives.

In this test the baseline transactional throughput was 30.3K batch requests per second and with PMM enabled the throughput increased to 41K batch requests per second. Use of Microsoft PMM on HPE NVDIMM technology increased the transactional throughput by 35% over regular SSD based log file.

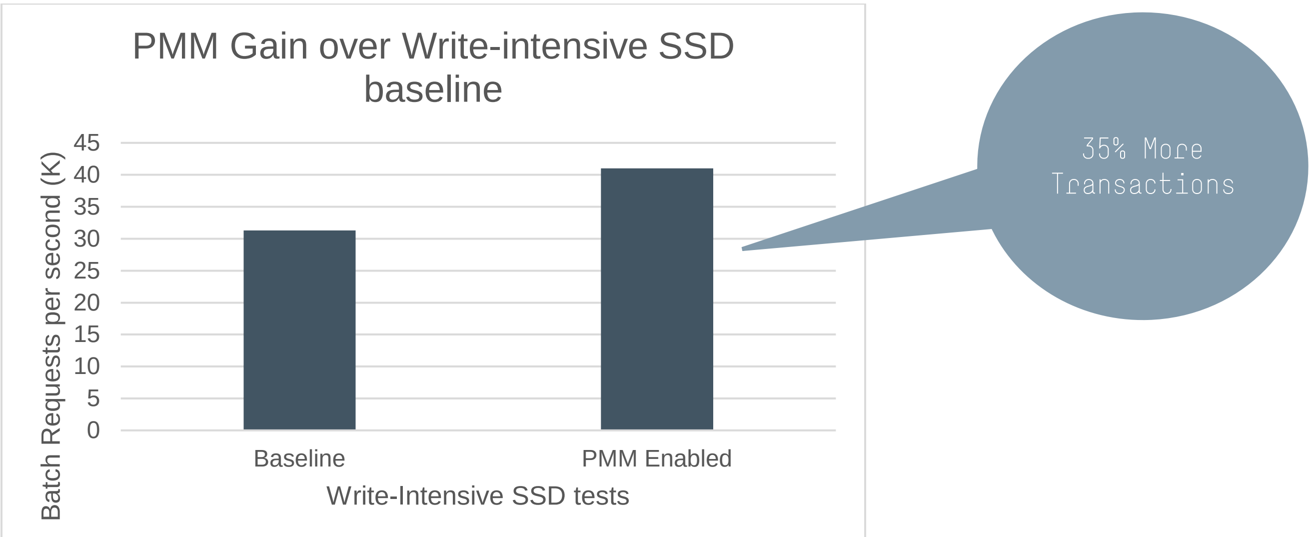

Figure 6. Transactional improvement over WI SSD baseline when PMM tail of the log is enabled

HPE PCIe Workload Accelerator cards

Figure 7 shows the transactional gain when PMM logging is enabled on a database deployed on HPE PCIe Workload Accelerators. HPE NVDIMM technology increases the transactional throughput by 9%.

In this test the baseline transactional throughput was 46.8K batch requests per second and with PMM enabled the throughput increased to 51K batch requests per second.

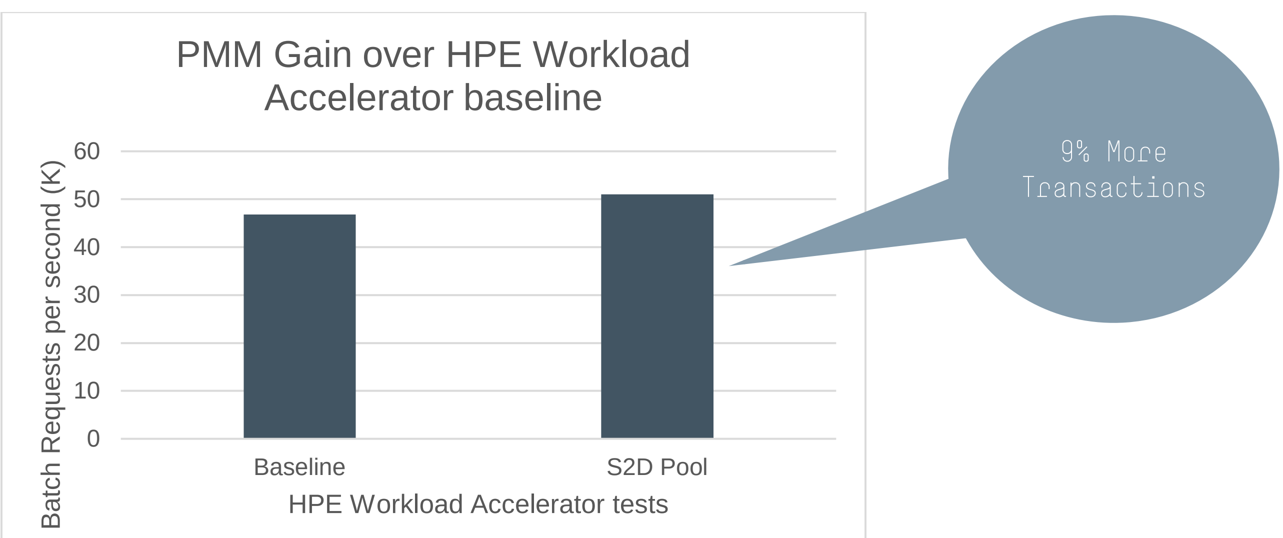

Figure 7. Transactional improvement over HPE PCIe Workload Accelerator baseline when PMM tail of the log is enabled

Analysis and recommendations

The testing has demonstrated the performance benefits of configuring NVDIMM modules for SQL Server 2016 tail of the log enhancements and when used as persistent media devices under Microsoft Storage Spaces.

A single NVDIMM is all that is needed for the current implementation of tail of the log. Multiple NVDIMMs can be combined into a Microsoft Storage Spaces pool to create a very low latency persistent media device for data storage and both techniques can be combined to have data residing on a Storage Spaces pool, and log file acceleration using tail of the log.

In addition, use of NVDIMMs as a cache tier in Storage Spaces retains high availability and resiliency by allowing mirroring between DIMM modules.

When both data and log acceleration approaches are used concurrently, further transactional throughput gains are achieved. This gain factor depends on the workload and its I/O characteristics.

Summary

The availability of HPE 8GB NVDIMM modules improves Microsoft SQL Server 2016 Database performance by combining the speed of memory access time with the data persistence of storage. The benefits potentially include license cost reductions, when databases are deployed on lower core count systems but accelerated via NVDIMM use to achieve par performance with its high core count legacy counterpart.

Implementing a proof-of-concept

As a matter of best practice for all deployments, HPE recommends implementing a proof-of-concept using a 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 HPE partner.

Appendix A: Bill of materials

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

Table 1. Bill of materials

QTY	PART NUMBER	DESCRIPTION
		HPE ProLiant DL380 Gen9 server
1	767032-B21	HPE DL380 Gen9 24SFF CTO Server
1	817947-L21	HPE DL380 Gen9 E5-2667v4 FIO Kit
1	817947-B21	HPE DL380 Gen9 E5-2667v4 Kit
8	728629-B21	HPE 32GB 2Rx4 PC4-2133P-R Kit
16	782692-B21	HPE 8GB NVDIMM Single Rank x4 DDR4-2133 Module
8	802582-B21	HPE 400GB 12G SAS WI 2.5in SC SSD
2	831735-B21	HPE SX350 1600GB PCIe Workload Accelerator
1	831733-B21	HPE SX350 1300GB PCIe Workload Accelerator
1	761872-B21	HPE Smart Array P440/4G FIO Controller
1	727250-B21	HPE 12Gb DL380 Gen9 SAS Expander Card
1	779799-B21	HPE Ethernet 10Gb 2P 546FLR-SFP+ Adptr
1	719073-B21	HPE DL380 Gen9 Secondary Riser
1	733660-B21	HPE 2U SFF Easy Install Rail Kit
2	720478-B21	HPE 500W FS Plat Ht Plg Pwr Supply Kit

Appendix B: Memory configuration

Table 2 shows the memory configuration for the HPE ProLiant DL380 Gen9 server, including which slots were populated with NVDIMMs and which ones had RDIMMs. For the memory population rules, see the HPE NVDIMM User Guide.

Table 2. Memory configuration

Memory Details (show empty sockets)

Memory Location	Socket	Status	HPE Memory	Part Number	Type	Size	Maximum Frequency	Minimum Voltage	Ranks	Technology
Processor 1	1	Good, In Use	HPE SmartMemory	752370-091	DIMM DDR4	32768 MB	2133 MHz	1.2 V	2	RDIMM
Processor 1	2	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 1	3	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 1	4	Good, In Use	HPE SmartMemory	752370-091	DIMM DDR4	32768 MB	2133 MHz	1.2 V	2	RDIMM
Processor 1	5	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 1	6	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 1	7	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 1	8	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 1	9	Good, In Use	HPE SmartMemory	752370-091	DIMM DDR4	32768 MB	2133 MHz	1.2 V	2	RDIMM
Processor 1	10	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 1	11	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 1	12	Good, In Use	HPE SmartMemory	752370-091	DIMM DDR4	32768 MB	2133 MHz	1.2 V	2	RDIMM
Processor 2	1	Good, In Use	HPE SmartMemory	752370-091	DIMM DDR4	32768 MB	2133 MHz	1.2 V	2	RDIMM
Processor 2	2	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 2	3	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 2	4	Good, In Use	HPE SmartMemory	752370-091	DIMM DDR4	32768 MB	2133 MHz	1.2 V	2	RDIMM
Processor 2	5	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 2	6	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 2	7	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 2	8	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 2	9	Good, In Use	HPE SmartMemory	752370-091	DIMM DDR4	32768 MB	2133 MHz	1.2 V	2	RDIMM
Processor 2	10	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 2	11	Good, In Use	HPE SmartMemory	N/A	DIMM DDR4	8192 MB	2133 MHz	1.2 V	1	R-NVDIMM
Processor 2	12	Good, In Use	HPE SmartMemory	752370-091	DIMM DDR4	32768 MB	2133 MHz	1.2 V	2	RDIMM

Resources and additional links

HPE Persistent Memory
hpe.com/servers/persistentmemory

HPE ProLiant DL380 Gen9
hpe.com/servers/dl380

HPE NVDIMM User guide
hpe.com/info/NVDIMM-docs

Storage Spaces Overview
https://technet.microsoft.com/en-us/library/hh831739(v=ws.11).aspx

Storage Spaces Direct in Technical Preview 4
https://blogs.technet.microsoft.com/clausjor/2015/11/19/storage-spaces-direct-in-technical-preview-4/

HPE Reference Architectures
hpe.com/info/ra

Intel Frequency Optimized CPU
http://ark.intel.com/products/92979/Intel-Xeon-Processor-E5-2667-v4-25M-Cache-3_20-GHz

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

Sign up for updates

★ Rate this document

Hewlett Packard Enterprise

4AA6-7123ENW, August 2016