

Accelerating Microsoft SQL Server Performance With NVDIMM-N on Dell EMC PowerEdge R740

A performance study with NVDIMM-N

Dell EMC Engineering
September 2017

Revisions

Date	Description
September 2017	Initial release

Copyright © 2017 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be the property of their respective owners. Published in the USA [9/21/2017] [White paper]

Table of contents

- Revisions.....2
- Executive summary.....4
- 1 NVDIMM-N architecture5
- 2 Microsoft enablement for storage class memory7
 - 2.1 SQL Server transaction commit processing8
- 3 Performance testing10
 - 3.1 Server components11
 - 3.2 Benchmark testing configuration11
- 4 Conclusion.....12
- 5 References13

Executive summary

OLTP databases are ubiquitous in enterprise applications including stock trading and order processing. These applications generally run short transactions with high frequency and require low latency. The capabilities of persistent storage are no longer able to keep up with the I/O demands and the latency requirement. The non-volatile dual in-line memory module (NVDIMM-N) technology enabled in the 14th Generation of Dell EMC PowerEdge servers has the potential to reduce latency and improve performance. NVDIMM-N is a new class of memory that provides storage persistence with the speed of memory. It combines NAND flash, DRAM, and an optional power source into a single memory subsystem. It delivers DRAM-like latencies and can back up any data it handles.

The Dell EMC PowerEdge R740 server is a powerful workload-optimized 2U, 2-socket platform capable of handling demanding workloads such as OLTP databases. The PowerEdge R740 supports NVDIMM-N capabilities.

For Microsoft SQL Server OLTP transactions, a significant performance bottleneck that causes latency issues occurs when committing to transaction log. SQL Server 2016 SP1 introduces a new capability that can accelerate transaction commit times by 2-4 times when using NVDIMM-Ns.

In this study, we compared the performance of PowerEdge R740 with NVDIMM-N and a similar system without NVDIMM-N. The results indicate that using NVDIMM-N:

- Increases transactions per second by ~42%
- Increases batch requests per second by ~43%

1 NVDIMM-N architecture

NVDIMM-Ns can be installed in a regular DIMM slot just like memory. At the same time, they have the ability to retain data in case of a power failure. The NVDIMM-Ns have onboard flash that are used to back up DRAM data and the data is retained in case of a power outage. A battery backup unit (BBU) is required to provide back-up power in case of AC power loss. The 14th generation of Dell EMC PowerEdge servers use host-managed NVDIMM-Ns with a single energy source to back up multiple NVDIMM-Ns in a server. Each server supports a single BBU that is cabled to the planar and shared across all NVDIMM-Ns.

NVDIMM-N comes in 8 GB or 16 GB capacity with a latency almost equal to DRAM. There are two access modes for NVDIMM-Ns—direct byte (DAX) or block oriented access.

NVDIMM-N brings storage persistence at the speed of memory by using an attached controller to transfer data from the volatile DRAM to the NAND Flash in the event of a power loss.

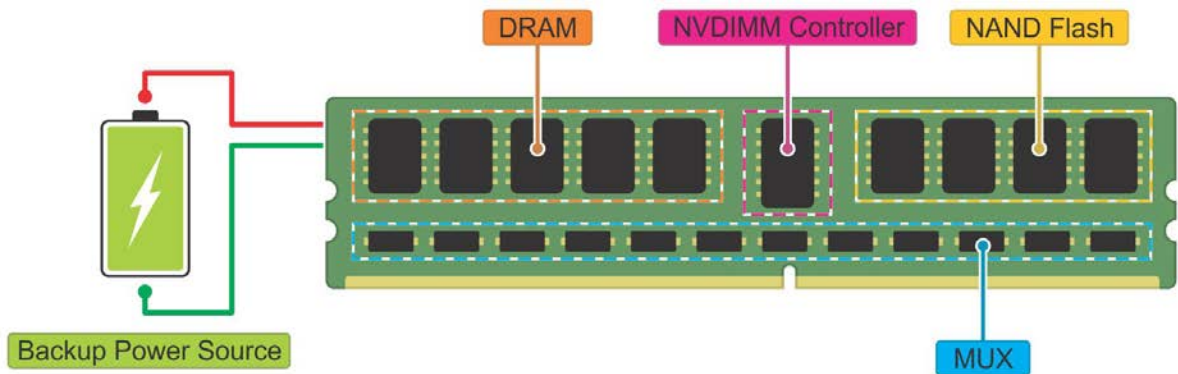


Figure 1 NVDIMM-N architecture

The installation guidelines for memory module in relation to NVDIMM-N must be followed. Below is an outline for PowerEdge R740:

- Each system can support memory configurations with 1, 2, 4, 6, or 12 NVDIMM-N
- Only RDIMMs are supported for use with NVDIMM-N
- Supported configurations have two processors and a minimum of 12x RDIMMs
- A maximum of 12 NVDIMM-Ns can be installed in a system

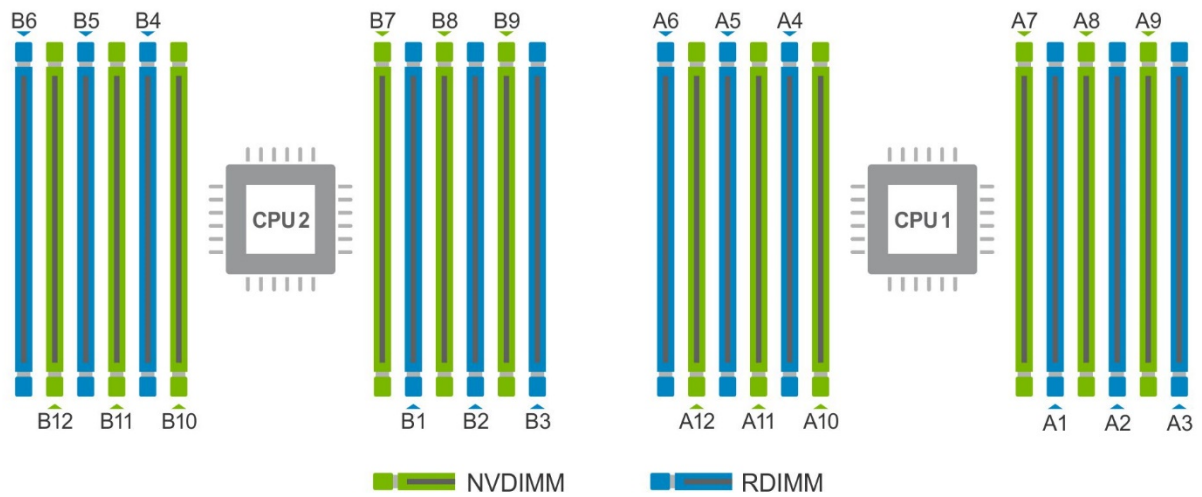


Figure 2 NVDIMM-N memory module in the system

2 Microsoft enablement for storage class memory

Windows Server 2016 recognizes NVDIMM-N modules as a disk device with special characteristics. It supports the following two access methods:

- Direct access (DAX) – DAX is a new protocol for storage access and is faster than traditional modes. If the file system is formatted as DAX, the operating system allows byte-level access to the persistent memory. DAX requires changes in the operating system.
- Block storage – Block storage is a traditional storage access protocol and is slower than DAX mode. Block storage does not require any changes in the operating system.

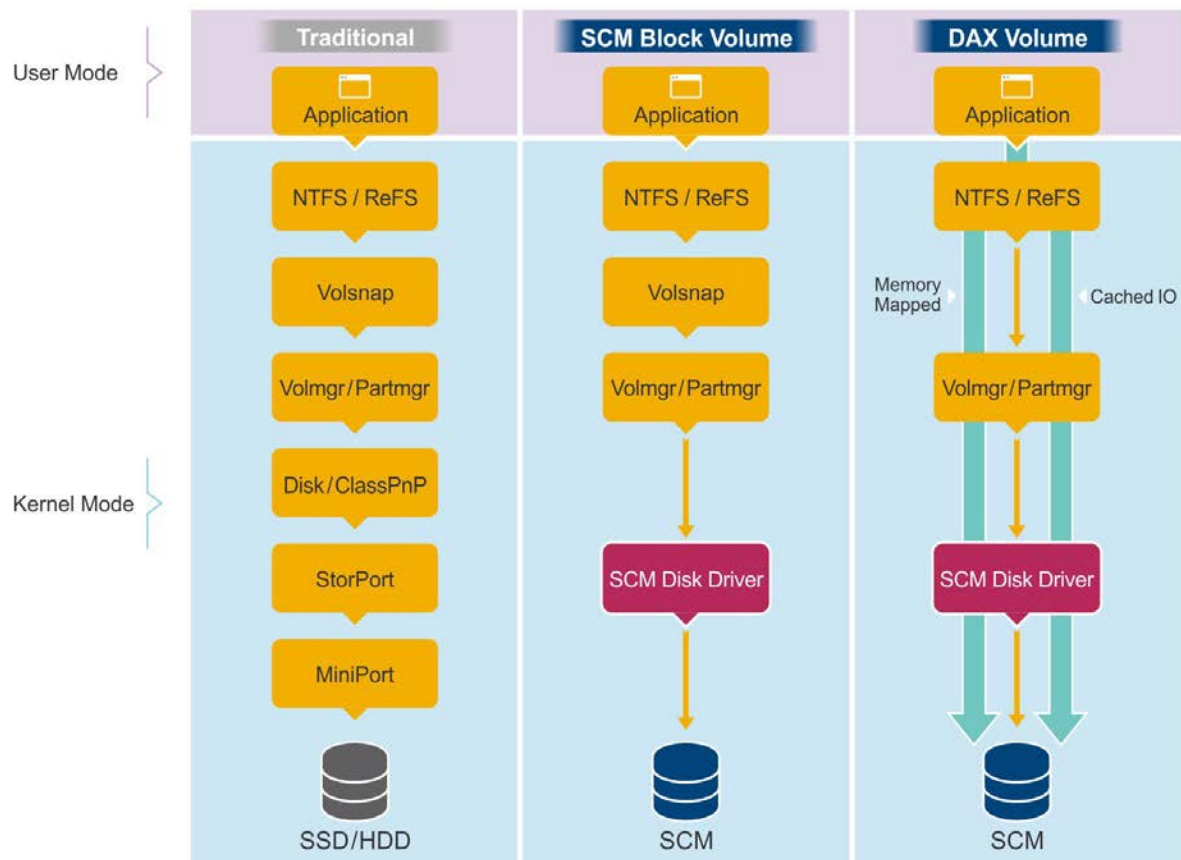


Figure 3 Storage class memory comparison

2.1 SQL Server transaction commit processing

SQL Server writes transaction log records in a buffer, and flushes them out to durable media during commit processing. SQL Server will not complete the commit until the commit log record is stored on a stable media. This log hardening can delay the processing in systems with a very high transaction rate.

Figure 4 shows how traditional log processing works. Each log record is copied in a buffer and is made to wait to be written to a disk. The log records are written to a disk in one of the two scenarios:

- The log buffer is full
- A commit record is written or sent

When either of these happen, all log records are transferred from the log buffer to disks in the transaction log file.

If a commit record comes in, the log records are written to disks even if the buffer is not full. If two commit records come in, the second record is made to wait until all log records are written to disks, and then the second commit request is executed.



Figure 4 Traditional log processing

As shown in Figure 5, SQL Server 2016 SP1 uses a region of memory that is mapped to a file on a DAX volume. This enables the memory to hold the log buffer. The memory hosted by the DAX volume is already persistent. Hence, SQL Server does not need to write the contents of the buffer to disks before executing another commit. As a result, commits are processed as soon as the transaction completes, without waiting for the log cache data to be written to disks.

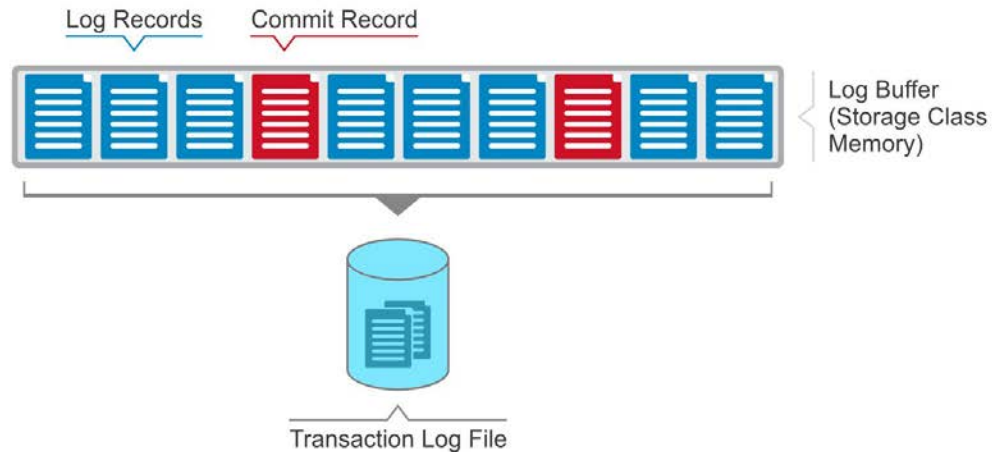


Figure 5 Log processing with SQL Server 2016 SP1

SQL Server Tail of Log must be enabled at the database level and assigned to NVDIMM-N DAX volume. If there are multiple databases, all the database tail of log can be assigned to a single NVDIMM-N DAX volume or they can be mapped to different NVDIMM-N DAX volumes.

In this study, four OLTP databases were tested with tail of log enabled on four different NVDIMM-N DAX volumes respectively.

3 Performance testing

A setup with four NVDIMM-N modules on Dell EMC PowerEdge R740 was utilized for testing. Using the TPC-C benchmark, the performance of four OLTP databases of 350 GB each was tested. SQL Server databases were deployed on 4 x 350 GB SSD drives. SQL server binaries and Temp DB files were deployed on 2 x 350 GB SSD drives. The performance was monitored using the following Windows performance counters for SQL Server:

- SQL statistics–Batch requests per second
- Databases–Transactions per second.

To determine the benefits of NVDIMM-N, two different scenarios were explored:

Scenario 1 (without NVDIMM-N)	Scenario 2 (with NVDIMM-N)
Tail of the log for all databases resided in volatile memory RDIMM	Tail of log for all databases resided in persistent memory NVDIMM-N. Tail of the log for the four databases were mapped to four different NVDIMM-N (DAX mode) modules respectively.

The following steps were performed for testing:

1. Deploying Microsoft SQL Server SP1
2. Creating four databases of 350 GB each in size
3. Using HammerDB, benchmarking a TPC-C workload for scenario 1 and 2 respectively
4. Running each benchmark test for 10 minutes with 16 users per database

As summarized in Figure 6, the test results indicate that there is a 42% increase in TPS and 43% increase in batch requests per second with the use of NVDIMM-N.

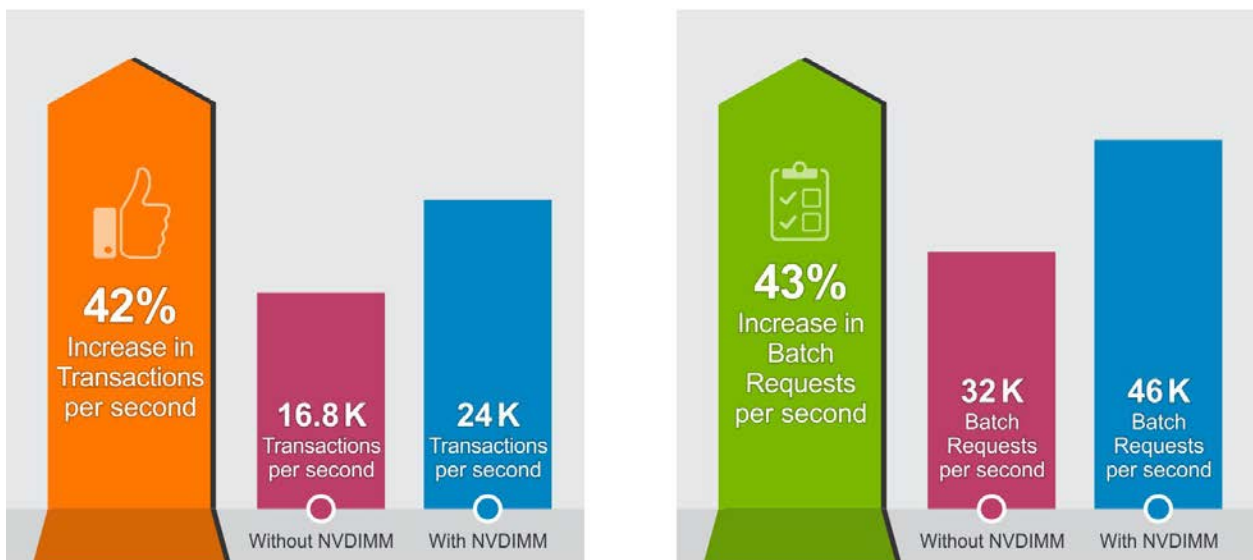


Figure 6 Performance comparison of SQL Server 2016 with NVDIMM-N and without NVDIMM-N

3.1 Server components

The setup that was tested included the following components:

Component	Specification (with NVDIMM-N)	Specification (without NVDIMM-N)
Server platform	Dell EMC PowerEdge R740	Dell EMC PowerEdge R740
Processor	2 x Intel Xeon Platinum 8176 @ 2.10 GHz with 28 cores	2 x Intel Xeon Platinum 8176 @ 2.10 GHz with 28 cores
Memory	384 GB @ 2666 MHz	384 GB @ 2666 MHz
NVDIMM-N	4 x 16 GB @ 2666 MHz	NA
Disks	4 x 1.6 TB SSD for data and log 2 x 1.6 TB SSD for temp DB	4 x 1.6 TB SSD for data and log 2 x 1.6 TB SSD for temp DB
Operating system	Windows Server 2016	Windows Server 2016
Database server	SQL Server 2016 SP1	SQL Server 2016 SP1

Note: For the purpose of this testing, we used a single server. For SQL Server running in a production environment, a minimum of two servers is recommended for ensuring high-availability (HA).

3.2 Benchmark testing configuration

The benchmark configuration used in the test is listed in the following table:

Setting	Value
Power plan	High Performance
SQL Server maximum memory	256 GB
MAXDOP	1
Lock pages in memory	Enabled for SQL Server service account

4

Conclusion

NVDIMM-N technology represents the latest development in persistent memory. NVDIMM-Ns combine storage persistence and memory speeds thus eliminating the time for data traversing through block storage stacks. The new PowerEdge R740 with the Intel Xeon Platinum processors configured with NVDIMM-Ns can process demanding workloads with unprecedented I/O requirements.

In this study, we evaluated the performance benefits of NVDIMM-N modules for a SQL Server 2016 OLTP workload running on the Dell EMC PowerEdge R740 servers with SQL Server 2016 SP1 tail of log caching. The study concludes that using NVDIMM-N in a DAX mode provides the following benefits:

- Increases transactions per second by ~42%
- Increases batch requests per second by ~43%

5 References

The following are referenced or recommended resources related to this study:

For more information on NVDIMM, see <https://www.micron.com/products/dram-modules/nvdimm#/>

For more information on Dell EMC PowerEdge R740 servers, see <http://www.dell.com/en-us/work/shop/povw/poweredge-r740>

For more information on Microsoft SQL Server Tail of the Log Caching, see <https://blogs.msdn.microsoft.com/bobsq/2016/11/08/how-it-works-it-just-runs-faster-non-volatile-memory-sql-server-tail-of-log-caching-on-nvdimm/>