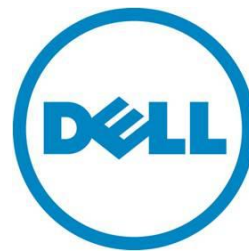

Achieving a High Performance OLTP Database using SQL Server[®] and Dell[™] PowerEdge[™] R720 with Internal PCIe SSD Storage

This Dell Technical White Paper discusses the OLTP performance benefit achieved on a SQL Server database using a Dell PowerEdge R720 server.

Jisha J

Dell Database Solutions Engineering

March 2012



This document is for informational purposes only and may contain typographical errors and technical inaccuracies. The content is provided as is, without express or implied warranties of any kind.

© 2012 Dell Inc. All rights reserved. Dell and its affiliates cannot be responsible for errors or omissions in typography or photography. Dell, the Dell logo, and PowerEdge are trademarks of Dell Inc. Intel and Xeon are registered trademarks of Intel Corporation in the U.S. and other countries. Microsoft, Windows, SQL Server and Windows Server are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and names or their products. Dell disclaims proprietary interest in the marks and names of others.

March 2012 | Rev 1.0

Contents

- Executive summary..... 5
- Introduction 5
- Dell PowerEdge R720 overview..... 5
 - Processors 7
 - Memory 7
 - Internal drive support with front-loading PCIe SSDs..... 7
- Database performance expectations from the PowerEdge R720..... 8
- Performance comparison with the previous generation 2U server 8
 - Test methodology..... 8
 - Test configuration 8
 - Results and analysis 10
- Server-based PCIe SSD drives and their impact on database performance..... 13
 - Test configuration 14
 - User database files on PCIe SSD drives..... 14
 - Temp database files and user database logs on PCIe SSD drives 18
- Conclusion 19
- Appendix A..... 21

Tables

- Table 1. PowerEdge 12th generation Customer Benefits over PowerEdge 11th generation..... 6
- Table 2. Memory population per channel and the memory operating speed for RDIMMs 7
- Table 3. Test configuration comparison of the PowerEdge R710 and PowerEdge R720 9
- Table 4. R720 vs. R710 Performance Comparison - Result Highlights..... 12
- Table 5. PowerEdge R720 with PCIe SSD test configuration details..... 14
- Table 6. SQLIO test parameters 21

Figures

- Figure 1. PowerEdge R720..... 5
- Figure 2. R720 vs. R710 performance comparison - transactions per second 10
- Figure 3. R720 vs. R710 performance comparison - response time 11
- Figure 4. R720 vs. R710 performance comparison - processor utilization 12
- Figure 5. PCIe SSD performance: transactions per second 15
- Figure 6. PCIe SSD performance: application response time..... 16
- Figure 7. PCIe SSD performance: CPU utilization 17
- Figure 8. PCIe SSD performance: average disk idle time 18

Achieving a High Performance OLTP Database using SQL Server® and Dell™ PowerEdge™ R720 with Internal PCIe SSD Storage

Figure 9. Database performance improvement achieved with PowerEdge R720 over the previous generation server..... 19

Figure 10. Database performance improvement achieved using PCIe SSD drives..... 20

Figure 11. SQLIO results comparison: OS RAID (mirrored disks) vs. no OS RAID for 4 * 350GB PCIe SSD drives 21

Executive summary

Dell™ has recently launched its next generation servers, including the PowerEdge R720 which is designed for performance. The support for the powerful processors, high memory capacity and the front loading PCIe SSDs makes it highly desirable for high performance demanding application deployments. Microsoft™ SQL Server® is one of the most popular databases which run many performance critical applications worldwide. SQL Server database applications may benefit by having this powerful server as the backend. This whitepaper explores the performance benefits of the PowerEdge™ R720 for a standard database workload. The CPU performance and the PCIe SSD performance are evaluated using specific OLTP (Online Transaction Processing) test configurations.

Introduction

Today, the industry expects high efficiency in all the data center products. The requirements for performance and power optimizations have also increased exponentially over previous years. The customer's desire for the best data center products drives enterprise giants like Dell to come up with better products every year.

In addition, the need for highly efficient hardware which enables high performing databases is increasing day by day. Currently, data center products are being designed with high workloads in mind to deliver high performance and operational efficiency. These highly efficient hardware products ensure the competent backend for the software components of the enterprise stack.

Dell recently announced its next generation of Enterprise servers which are aimed at optimizing overall datacenter efficiency. The intent of this whitepaper is to highlight the performance advantages of using the latest 2U server from Dell for a SQL Server database environment.

Dell PowerEdge R720 overview

The Dell PowerEdge R720 is a mainstream 2S 2U rack server focused on performance and scalability. It is a powerful server offering extreme storage performance with PCIe SSD cards and supports high bandwidth network I/O with 10GbE support.

Figure 1 shows the front view of a PowerEdge R720 server.

Figure 1. PowerEdge R720



Achieving a High Performance OLTP Database using SQL Server® and Dell™ PowerEdge™ R720 with Internal PCIe SSD Storage

The major technology enhancements of the PowerEdge R720 compared to the previous generation servers is given in Table 1.

Table 1. PowerEdge 12th generation Customer Benefits over PowerEdge 11th generation

12G PowerEdge R720	Customer benefit over 11G
New Intel® Xeon® Processor E5-2600 Family	More efficient, faster compute results
iDRAC7 with Lifecycle Controller Agent-free monitoring Auto-update for replacement parts	Maximize uptime Monitor a server even if it goes off-line Reduce maintenance time
Network Daughter Card (NDC) & Select Network Adapter (SNA) Network fabric tailored to customer needs Switch-independent partitioning	No vendor lock-in Investment protection: upgrade when you are ready Flexibly allocate resources among VMs
Front-loading, hot-swappable PCIe SSD drives	Turn data into insights faster; up to 3x performance of traditional SSD ¹
PowerEdge RAID Controller (PERC) Industry-leading IOPS with H710P/H810 Split mirror	Performance improvement over the earlier generation PERC Replace drives quickly
Right-sized PSU	Save costs with energy efficiency
Power monitoring (optional capping)	Control cooling costs
R720 internal GPU	Performance improvement in Virtual Desktop Infrastructure (VDI)/ High-Performance Computing (HPC)
Memory Maximum density; more flexibility in building configurations and enabling RAS features	No compromise on virtualization Easier ordering/configuration

¹ PCIe SSD Storage, SNIA, by Gary Kotzur, Dell Inc., snia.org/sites/default/education/tutorials/2011/fall/SolidState/GaryKotzur_PCIe_SSD_Storage.pdf.

The major components/features of the server are explained in the sections below.

Processors

The Dell PowerEdge R720 supports the latest Intel Xeon® Processor E5-2600 Family (Sandy Bridge-EP architecture) processors, which offer a wide range of features. One of those features is the QuickPath Interconnect (QPI) links which offer high-bandwidth inter-socket communication. It supports up to eight execution cores per processor and each core supports up to 2 threads for up to 16 threads per processor (using Hyper-Threading). These processors supports up to 20MB of Last Level Cache (LLC) shared among all the cores (up to 2.5MB per core).

Memory

The Dell PowerEdge R720 supports DDR3 DIMMs ((Un-buffered DIMMs (UDIMM ECC), Registered DIMMs (RDIMM), and Load Reduced DIMMs (LRDIMM))) with ECC. Single, Dual and Quad rank DIMMs are supported. 2 voltages are supported: Standard DIMM voltage of 1.5v which support speeds up to 1600, and Low Voltage (1.35 V) which supports speeds up to 1333Mhz. Up to 12 DIMMs can be installed per CPU and up to 24 DIMMs can be installed in a dual-socket configuration. R720 has four memory channels per CPU and each channel will support up to 3 DIMMs. As of now, it supports a maximum memory capacity of 768GB.

Operational speeds of 1600, 1333, 1066 and 800 are supported, depending on the DIMM type, DIMM capability and DIMMs per channel populated. Table 2 shows the memory DIMM population and the memory operating speeds for 1.5v operating voltage.

Table 2. Memory population per channel and the memory operating speed for RDIMMs

DIMMs Populated per channel	Memory Operating Speed for 1.5v operating voltage	Max Ranks for DIMM in Channel
1	1600, 1333, 1066 (Quad Rank 1066,800)	Quad Rank
2	1600,1333, 1066 (Quad Rank 800)	Quad Rank
3	1066, 800	Dual Rank

Internal drive support with front-loading PCIe SSDs

The R720 supports up to 8*3.5” SAS, SATA and Near-Line SAS drives without an optional flex bay. With an optional flex bay, it supports up to 16 * 2.5” SAS, SATA, Near-Line SAS drives and PCIe SSD drives. The support for the PCIe SSD drives enables the PowerEdge R720 to deliver high storage IOPS internally. The entire range of 8th generation PERC (PowerEdge RAID Controller) controllers provides the hardware RAID capabilities and many other RAID options to ensure a reliable configuration. The PERC H710P and H810 controllers support high IOPS (Input Output Operations per Second), with SAS/SATA disks, to deliver premium in-box performance.

Achieving a High Performance OLTP Database using SQL Server® and Dell™ PowerEdge™ R720 with Internal PCIe SSD Storage

More information on the PCIe SSD drives and the performance improvement observed on a standard database workload are explained in the latter section of the whitepaper, [Server-based PCIe SSD drives and their impact on database performance](#).

For the complete feature details of the PowerEdge R720, please refer to Dell.com/PowerEdge.

Database performance expectations from the PowerEdge R720

The 12th generation PowerEdge servers come with the powerful components which raises application performance expectations. Database applications are one of the application categories where performance plays a very critical role. Different database applications have specific performance requirements to meet the client satisfaction. For example, OLTP (Online Transaction Processing) applications benefit from better TPS (Transactions per Second) and response time, whereas OLAP (Online Analytical Processing) demands more database throughput. The PowerEdge R720 has high processor bandwidth, large memory capacity and huge storage performance capacity, which makes it ideal for database applications.

This whitepaper intends to evaluate the database performance that the PowerEdge R720 can deliver, using the standard SQL Server 2008 R2 database configuration. The scope of this whitepaper is limited to the OLTP workload. Other workloads including OLAP/DSS, exchange etc. are not covered in this whitepaper.

The latter sections of this whitepaper focus on the performance impact of using the PowerEdge R720 for OLTP database workloads.

Performance comparison with the previous generation 2U server

To see the performance improvement that the PowerEdge R720 brings, compared to the earlier generation PowerEdge R710, we conducted experiments at Dell Labs to compare the performance of the PowerEdge R710 and R720. The latter sections of this whitepaper explain the test methodology, configurations details and observations of the comparison study.

Test methodology

The tests were designed in such a manner to evaluate the processor capabilities of the PowerEdge R710 and R720. An OLTP test tool was used to simulate a TPC-E stress workload from a client machine. The same test benchmark was executed on both the PowerEdge R710 and PowerEdge R720.

Test configuration

The hardware and software component details are listed in Table 3. The PowerEdge R710 was populated with the highest performing processor supported and the maximum memory possible. This helped us in getting the CPU stressed to the maximum extend, without being bottlenecked much at the disk backend. The memory capacity required for testing the PowerEdge R720 was decided based on the memory population guidelines to achieve the maximum memory speed of 1600MHz. The memory populated was 128GB, which is less than the R710 memory of 144GB. Anyhow, it was verified that the workload was not posing much of memory or disk bottleneck for R720.

Table 3. Test configuration comparison of the PowerEdge R710 and PowerEdge R720

Component	PowerEdge R720	PowerEdge R710
Server BIOS	0.3.35	6.1.0
Memory(RAM)	128GB (8 * 16GB DDR3 1600MHz DIMMs)	144GB (18 * 8GB DDR3 1333MHz DIMMs) - Maximum Memory Supported
Processor	2 * 8C Intel Xeon 2690 (Sandy Bridge) processor @ 2.7GHz (Hyperthreading Enabled)	2 * 6C Intel Westmere X5690 processor @ 3.47GHz (Hyperthreading Enabled)
Host Bus Adapters	2 * Qlogic 8Gbps QLE2562 Dual Port HBAs	2 * Qlogic 8Gbps QLE2562 Dual Port HBAs
External Storage	1 * PowerVault MD3620F(24 * 15k SAS 6Gbps 146GB disks)	1 * PowerVault MD3620F(24 * 15k SAS 6Gbps 146GB disks)
Operating System	Microsoft Windows 2008 R2 SP1	Microsoft Windows 2008 R2 SP1
Database	Microsoft SQL Server 2008 R2 SP1	Microsoft SQL Server 2008 R2 SP1
Database Size	500GB	500GB
Database Recovery Model	Full	Full
Test Workload	TPC-E Benchmark(OLTP)	TPC-E Benchmark(OLTP)

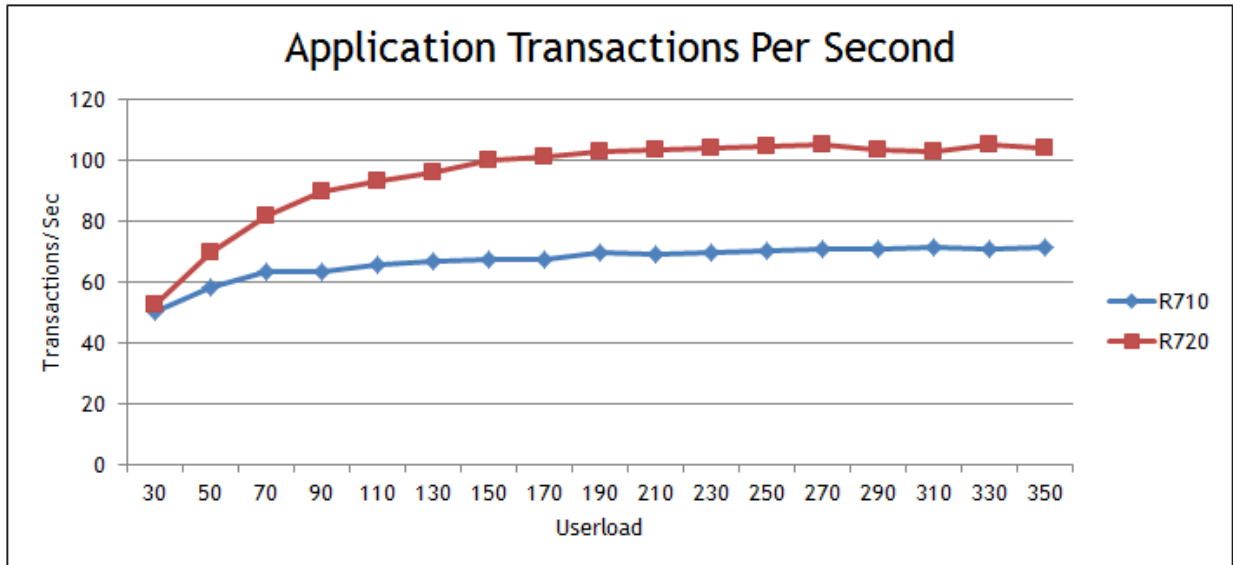
The test database of size 500GB was configured in the external storage. The database consisted of two datafiles on two separate logical drives (two LUNs each of 10 disks on RAID 10) and one log file on a separate RAID 1 LUN. The tempdb files were also distributed among the two data LUNs.

The test workload selected is TPC-E, because of the CPU-intensive nature of the transactions involved. Note: TPC-E is a very different and more IO and CPU intensive workload compared to TPC-C. Therefore, the performance numbers should not be compared against any TPC-C workload.

Results and analysis

The stress tests conducted showed that the PowerEdge R720 outperformed the R710 for the specific TPC-E workload. Figure 2 shows the Transactions per Second (TPS) delivered by R720 in comparison with the R710.

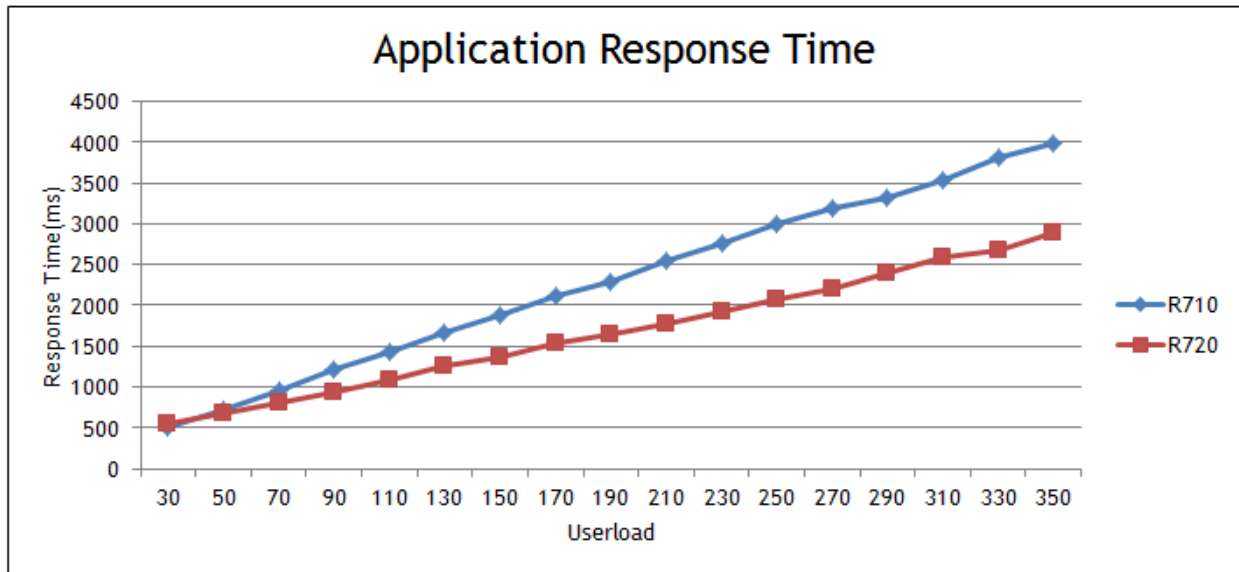
Figure 2. R720 vs. R710 performance comparison - transactions per second



The workload showed 1.5 times better TPS with the R720 compared with R710. In addition, it's very apparent from the graph that the R720 is able to consistently push a TPS of 105, even at high stress.

Figure 3 represents the application response time observed during the workload.

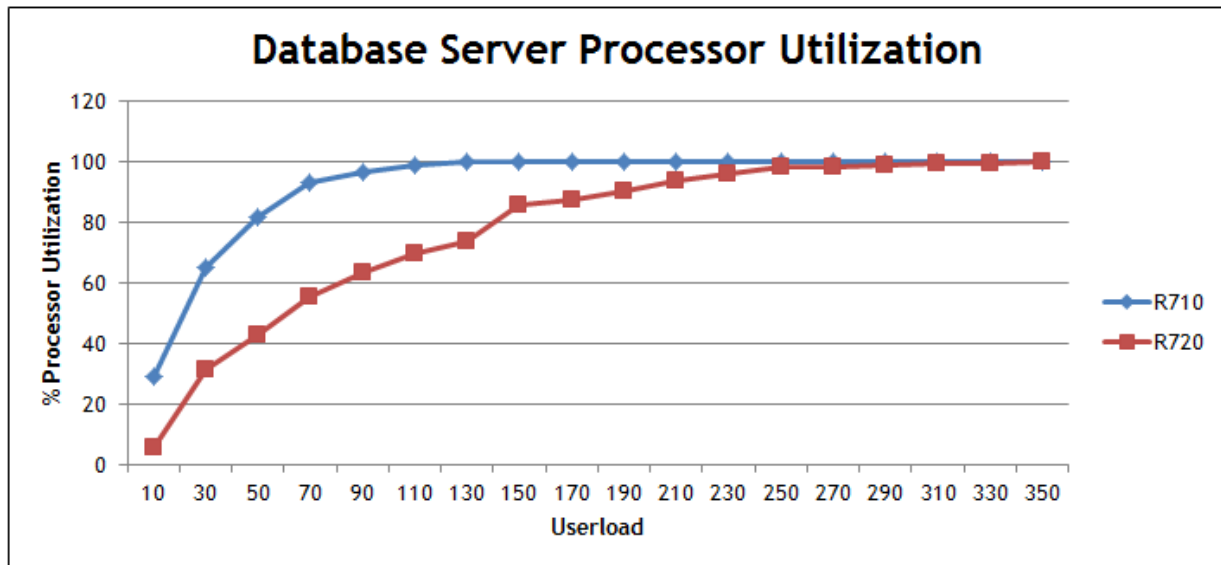
Figure 3. R720 vs. R710 performance comparison - response time



The response time numbers show that the R720 is able to deliver high TPS with a quicker response time, compared to the R710. The application response time showed up to 30% improvement with the R720 compared to a similar R710 configuration. As mentioned earlier, the R710 was configured with the maximum memory capacity possible (144GB) and the R720 was populated with 128GB of RAM, running at 1600MHz (the maximum memory speed supported by R720).

Figure 4 shows the processor utilization for the scenarios during the exercise. The processor utilization was observed to reach 100% utilized beginning with 110 users in the R710, whereas in the R720 the processor utilization did not reach 100% until 250 users.

Figure 4. R720 vs. R710 performance comparison - processor utilization



The other major highlights of the test results are listed in Table 4.

Table 4. R720 vs. R710 Performance Comparison - Result Highlights

Performance Variable	PowerEdge R720 Configuration			PowerEdge R710 Configuration		
	Min	Max	Average	Min	Max	Average
Bytes/Second(KBps)	102	203	183	96	138	128
Rows/Second	2738	5476	4943	2577	3718	3466

On average, the PowerEdge R720 delivered around 43% more Bytes per Second and Rows per Second for the specific database workload.

To summarize, the PowerEdge R720 features the latest best performing processor support and high speed memory support, which enables it to outperform the R710 for the specified database workload.

Server-based PCIe SSD drives and their impact on database performance

SSD(Solid State Disk) drives use NAND Flash memory chips to store the data, instead of the mechanical media used in traditional hard drives. They promise very low latency for disk access, due to the absence of any moving part. Traditional SSDs have used the interfaces designed to support mechanical drives. The current PCIe-based SSDs²³ make use of specialized drivers to communicate over the high speed PCIe interface to connect to the multiple channels of memory. This helps in providing low latency access compared to the quickest traditional storage arrays. PCIe SSDs would be very beneficial to the database applications that demand very low latency or that generate huge amounts of random read-write operations.

The PowerEdge R720 promises high storage throughput due to the support for the high performance PCIe SSD drives. The R720 supports up to 4 PCIe SSD drives in the front drive bay, in addition to SAS/SATA near-line SSDs. The front-loading PCIe SSD drives offer more flexibility (hot-plug support) and price benefit over other PCIe SSD vendors in the market.

Using SQLIO, we observed that the 4* 350GB PCIe cards were able to deliver a Read IOPS of over 600,000 8K random Input/output operations per second (IOPS) and over 400,000 8K random write IOPS.

We conducted a series of experiments to analyze how the high I/O based database workload would benefit by using PCIe SSD drives. For this exercise, we selected a lesser memory configuration compared to the earlier R710 comparison test. The SQL Server target memory was restricted to 8GB to direct more IOs to the disk backend by reducing the database buffering.

As of now, the PCIe SSDs support only software RAID. We recommend not using software RAID on the PCIe SSD cards, as performance degradation is being observed on the drives at high stress. Refer to Appendix A for more information. For high availability, it is recommended to have an application level high availability solution like database mirroring to be configured for the system.

The latter portions of this whitepaper discuss the performance impact of having the various database structures on the PCIe SSD tier.

² <http://searchsolidstatestorage.techtarget.com/tip/Is-PCIe-SSD-right-for-you-Deploying-PCI-Express-SSD-devices>

³ <http://searchstorage.techtarget.co.uk/feature/PCIe-SSD-What-it-is-and-how-you-can-use-it>

Test configuration

The system test configuration chosen for the PCIe SSD performance tests is shown in Table 5.

Table 5. PowerEdge R720 with PCIe SSD test configuration details

Component	PowerEdge R720 Details
Server BIOS	0.3.35
Memory(RAM)	16GB @1600MHz(2 * 8GB DDR3 DIMMs) [SQL Memory Restricted to 8GB, to direct more IOs to disk]
Processor	2 * 8C Intel Xeon 2690 processor @ 2.7GHz (Hyperthreading Enabled)
Host Bus Adapters	2 * Qlogic 8Gbps QLE2562 Dual Port HBAs
External Storage	1 * PowerVault MD3620F (24 * 15k SAS 6Gbps 146GB disks)
Internal PCIe SSD	4 * 350GB Micron Real SSD drives; Firmware - B1390008; Driver - 6.24.0.8
Operating System	Microsoft Windows 2008 R2 SP1
Database	Microsoft SQL Server 2008 R2 SP1
Database Size	500GB
Database Recovery Model	Full
Test Workload	TPC-E Benchmark(OLTP)

A very thorough knowledge of the application and I/O behavior is needed to optimally use the performance capability of the PCIe SSD drives. Careful examination of the workload needs to be conducted to determine the best candidate for PCIe SSD storage. For the current exercise, we examined the performance impact by placing the following database components on the PCIe SSD tier. The same TPC-E database workload was executed and the results were compared to the 15k SAS backend configuration (MD3620F with 24 disks).

User database files on PCIe SSD drives

The data files for a user database are one of the candidates to be placed in the SSD tier. It is advisable to have the small databases placed in the PCIe SSDs in their entirety. For larger databases, the most I/O intensive components of the database should be identified and moved to the SSDs for achieving high application performance. The query behavior also plays a major role in determining the proper candidates for SSD placement.

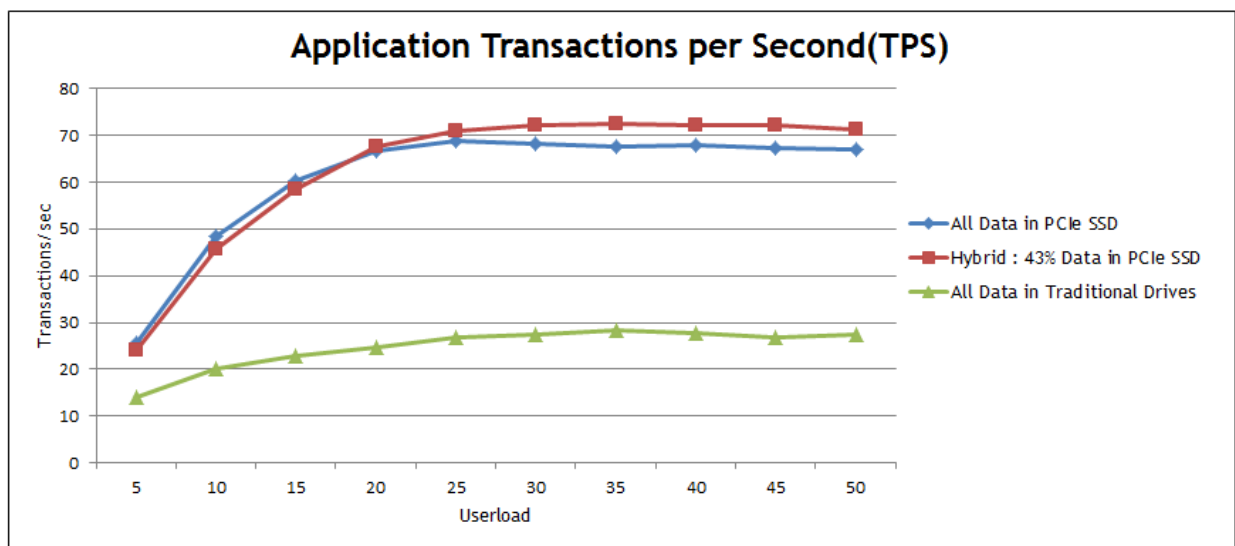
Achieving a High Performance OLTP Database using SQL Server® and Dell™ PowerEdge™ R720 with Internal PCIe SSD Storage

During the whitepaper exercise, the following database components were included on the PCIe SSD tier:

- **All the user data files** were moved to the SSD tier. (The database logs and the tempdb files were placed on traditional storage.)
- **The most accessed data subset** was determined by analyzing the query behavior. The Trade table was found to be the most I/O intensive structure of the workload. This data subset was moved to the PCIe SSD tier to analyze the benefit of these drives. The whole table was around 43 % (around 200GB) of the entire database size.

Figure 5 shows the improvement in TPS achieved due to the PCIe SSD backend.

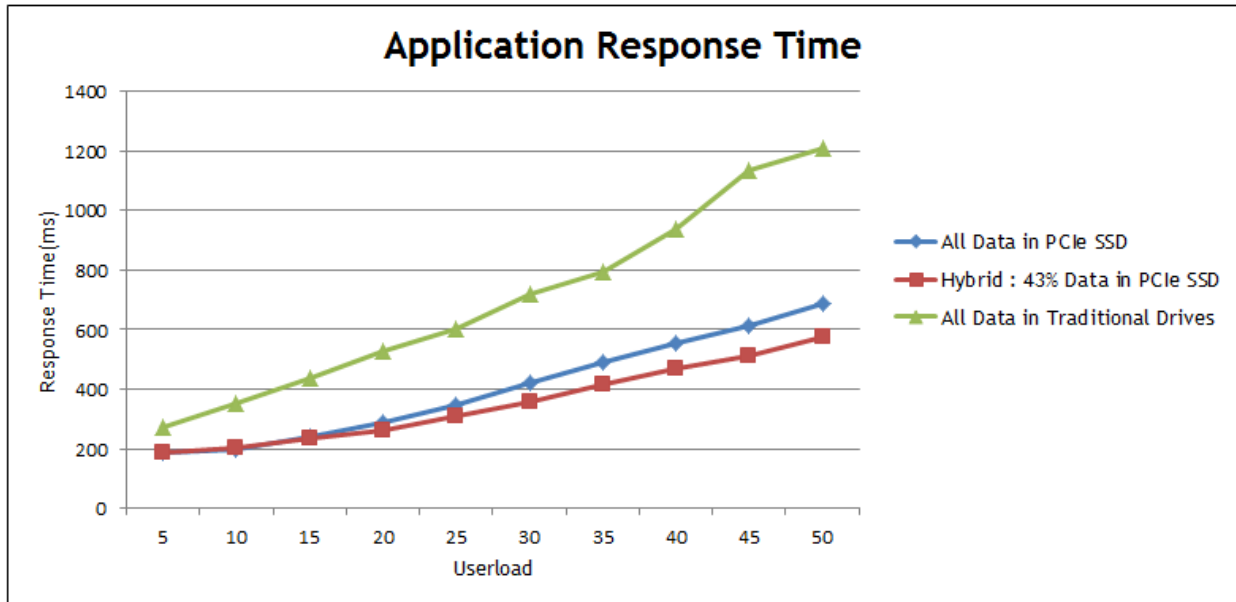
Figure 5. PCIe SSD performance: transactions per second



We observed up to 2.7 times improvement in TPS when we had the PCIe SSD layer introduced to host the database files. The hybrid use case where we had the highly accessed data in the PCIe SSDs was observed to deliver even better TPS than when we had the complete database on the SSD drives. This is probably due to the fact that in the hybrid case, the TPS is contributed from both the PCIe SSDs and the traditional Data LUNs. In addition, the PCIe SSD tiering also helped relieve the stress on the peer traditional data LUNs (also in the external storage) which helped improve the TPS numbers.

The same trend is visible in the case of the application response time as well. Figure 6 shows the application response time observed during the exercise.

Figure 6. PCIe SSD performance: application response time



We can see that the PCIe SSD layer helped to improve the overall application response time by up to 120% better than with the traditional drives.

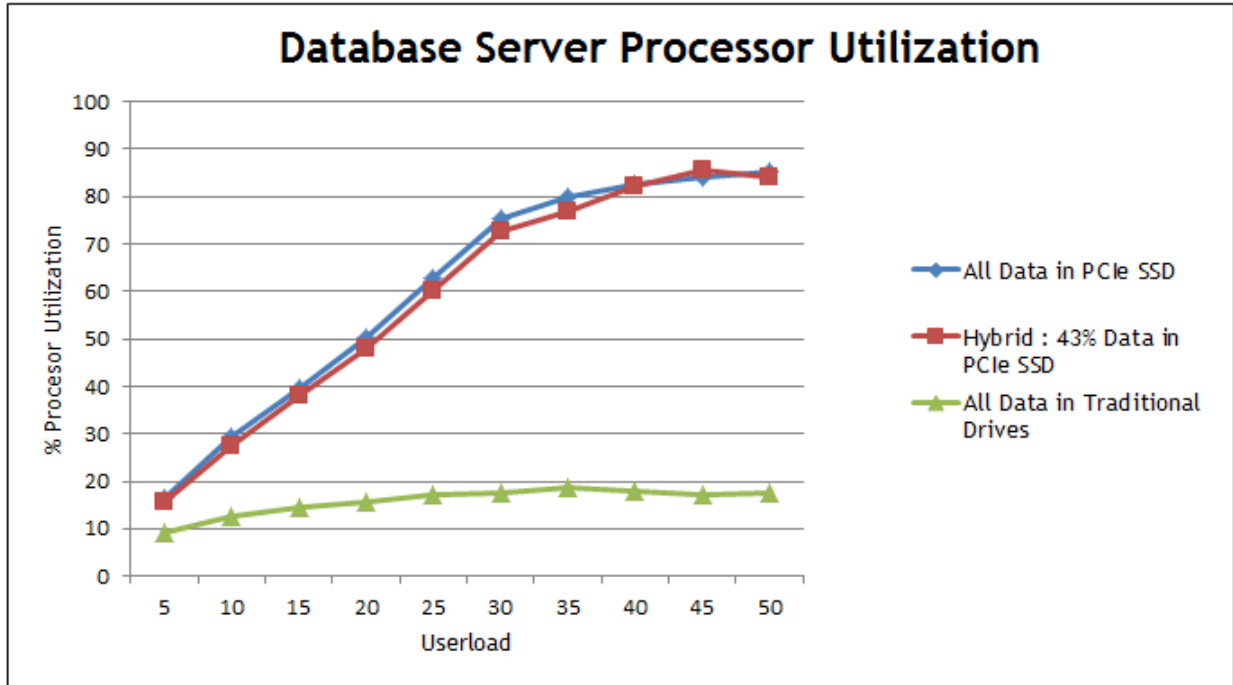
Another important trend to notice is that the TPS tends to flatten out beginning with 20 users when all of the data is placed in the SSD drives. In the hybrid scenario we observe the TPS flattening from 30 users. Whereas, when data is in the traditional storage, the TPS flattens out from around 20 users.

The reason for this is that the nature of the workload is CPU-intensive as well as disk I/O intensive. With up to 20 users, the workload is more disk intensive rather than cpu intensive. With more than 20 users, due to the database buffering, the read-intensive transactions shift the bottleneck towards the CPU.

Beyond this data point, the TPS flattens out due to the building CPU pressure, as well as the I/O pressure on the server. Here, the overall TPS of the system is bottlenecked at the earlier stages at the disk and later the performance benefit shifts both to the CPU and disk. Since the PCIe SSD drives alleviate the disk bottleneck efficiently, a less CPU-intensive workload would achieve maximum benefit by the PCIe SSD performance.

Figure 7 shows the CPU utilization for these cases.

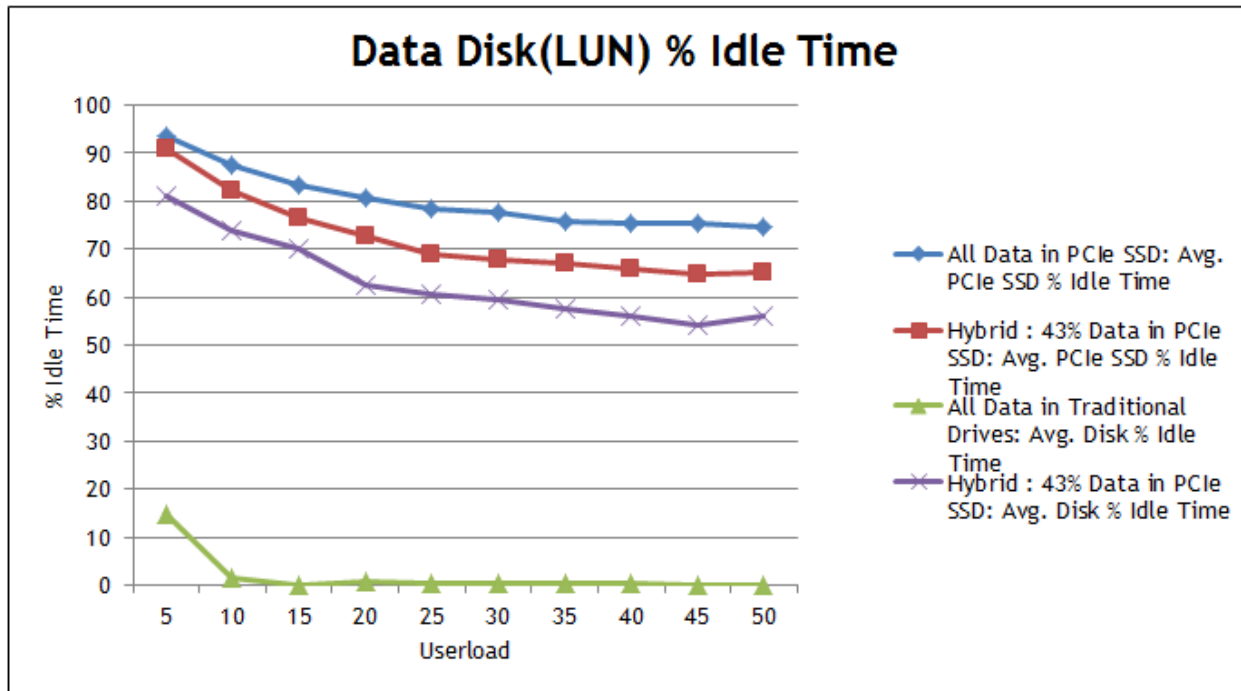
Figure 7. PCIe SSD performance: CPU utilization



From Figure 7, it is very obvious that the lower CPU utilization is due to the disk bottleneck by the workload. PCIe SSD drives relieve the bottleneck on the disk due to its high IOPs capability, which builds up more pressure on the CPU.

Figure 8 represents the data disk idle time during the exercise.

Figure 8. PCIe SSD performance: average disk idle time



When all the data is in the traditional disks, the disk utilization is almost 100%. In the Hybrid use case the 30 % use of the PCIe SSD drives made the overall disk busy time drop from 100% to 38%. This indicates the high speed data request processing capability of PCIe SSD drives.

Temp database files and user database logs on PCIe SSD drives

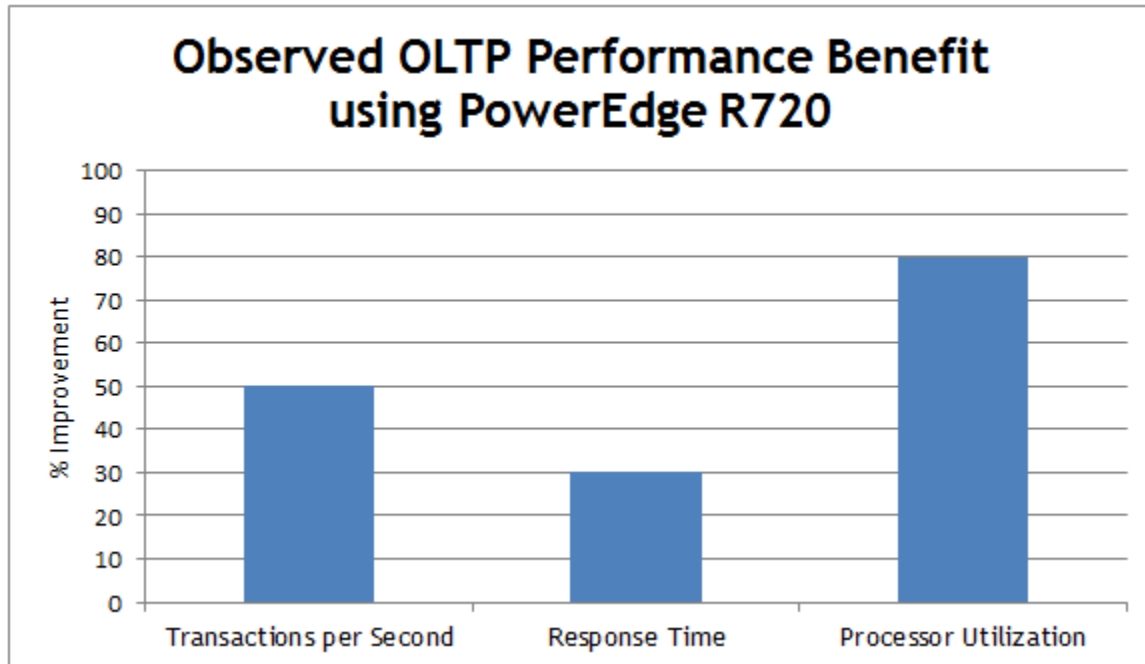
For the specific workload studied (read intensive), there was less pressure on the Temp database and the user database log files, compared to the data file disks. Therefore, we could not see any performance improvement by having only the Temp DB and log files on the PCIe SSDs. The workloads characterized by heavy inserts, updates, sorts etc. may be benefitted by the PCIe SSD performance.

It should be noted that SSDs may be used to provide a high performance disk backend for an application, but the maximal performance capability of the SSDs may not be realized in its entirety if the application is bottlenecked at the other database layers such as the CPU, network, or other data disks. Therefore, a thorough analysis of the current environment is recommended prior to the PCIe SSD installation to ensure the true PCIe SSD performance benefit.

Conclusion

The database performance experiments showed significant performance impact by using the PowerEdge R720 compared to the previous generation R710 server. The observed maximum database performance numbers for the PowerEdge R720, for the specific TPC-E database workload, are summarized in Figure 9.

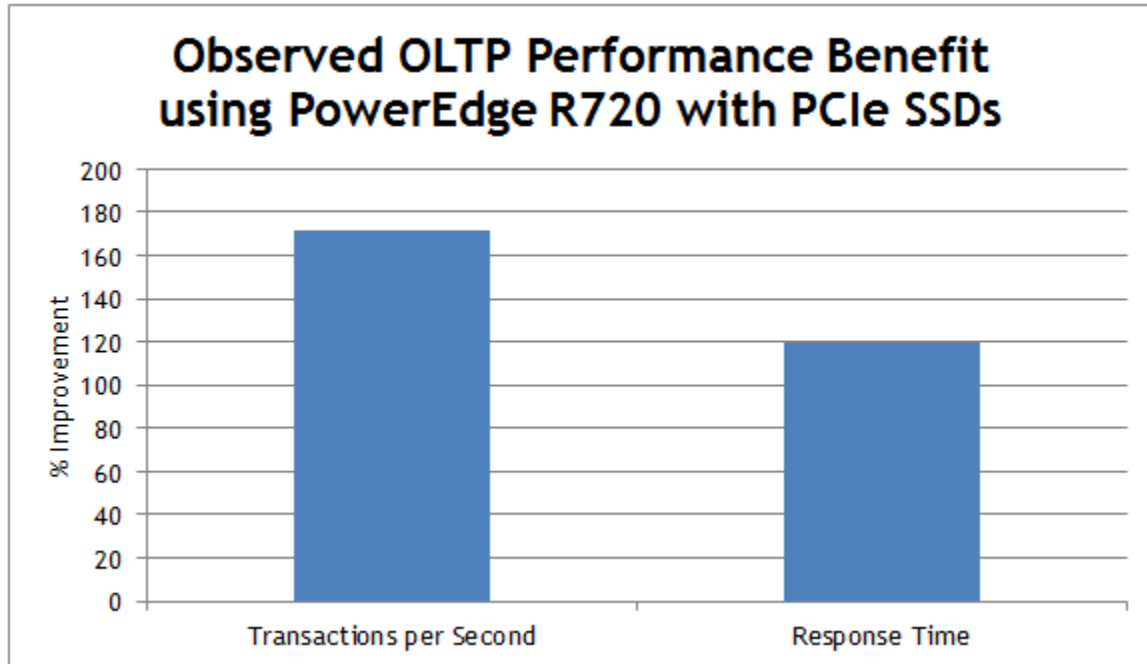
Figure 9. Database performance improvement achieved with PowerEdge R720 over the previous generation server



Achieving a High Performance OLTP Database using SQL Server® and Dell™ PowerEdge™ R720 with Internal PCIe SSD Storage

The high performing PCIe SSD drives supported by the PowerEdge R720 may be extremely helpful in relieving the disk bottleneck experienced by the databases. In particular, the PCIe SSDs may be highly beneficial for OLTP workloads in providing high IOPS (Input Output Operations per Second) and very low latency. The observed performance impact on using PCIe SSD drives for storing the database structures is summarized in Figure 10.

Figure 10. Database performance improvement achieved using PCIe SSD drives



To summarize, the PowerEdge R720 is a powerful server with high database processing capability due to powerful CPU support, high memory capacity and high-speed DIMMs. The high-speed front-loading PCIe SSDs help support dramatically high performing database applications efficiently. The major advantages of using the PowerEdge R720 (with PCIe SSDs) as a database server would be:

- Better Database Transactions per Second
- Better Database Response Time

Appendix A

Figure 11. SQLIO results comparison: OS RAID (mirrored disks) vs. no OS RAID for 4 * 350GB PCIe SSD drives

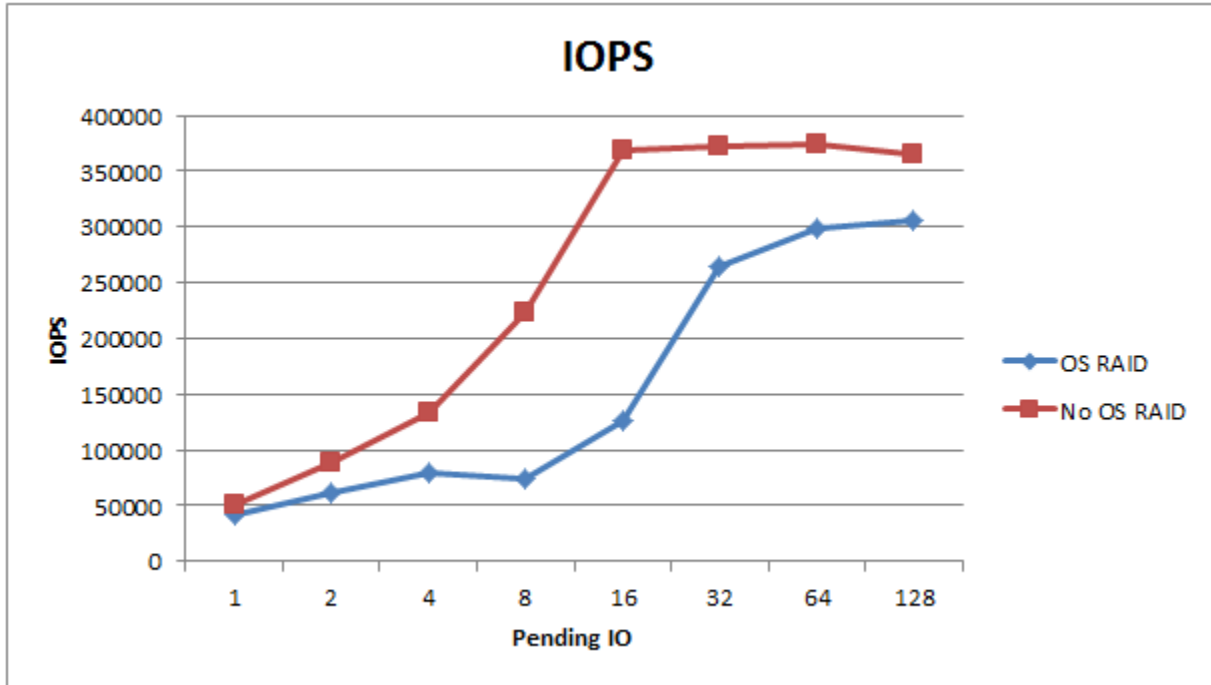


Table 6. SQLIO test parameters

Threads	4
IOType	Random
FileSize	51200 MB
Duration	30