Accelerating SQL Server 2014 OLTP Workloads on Hyper-V with SC4020 and FX Converged Architecture

Dell Storage Engineering
August 2015

A Dell Technical White Paper

Revisions

Date	Description
August 2015	Initial release

Acknowledgements

Dell Storage Applications Engineering: Mike Matthews, SQL Server Product Specialist

Dell Storage Solutions Marketing: Omar Rawashdeh

Editing: Daniel White

Table of contents

Executive summary

Many IT departments are challenged to address performance issues caused by the changing demands of online transaction processing (OLTP) applications. The workload produced by a given application can increase dramatically over time, due to an increasing number of users, database growth, or both. Database and application servers are often virtualized, which provides flexibility and simplifies hardware upgrades to address performance problems. By leveraging virtualization on Dell servers and storage hardware, OLTP workloads can be easily accelerated to meet those increasing workload demands.

This white paper examines a Microsoft® SQL Server® 2014 OLTP workload, virtualized using Windows Server® 2012 R2 Hyper-V®. The workload is generated by four application servers, each configured to simulate 500 concurrent users. Each application server connects to its own database server, each hosting a 520GB OLTP database. In total, the workload simulates 2000 users accessing over 2TB of data. The workload was run once on the baseline configuration and again after adding additional hardware.

The baseline configuration, representing an environment where workload demands have exceeded available hardware resources, consists of two Dell PowerEdge™ FC630 servers in a PowerEdge FX2 chassis and forty-eight 15K spinning drives in a Dell Storage SC4020 storage array. To create the accelerated configuration, the baseline configuration was upgraded by adding two FC630 servers to the FX2 chassis and solid-state drives (SSDs) were used instead of 15K drives in the SC4020.

The hardware upgrades gave the OLTP workload a very large performance boost, greatly increasing the number of transactions performed in the database and significantly reducing response time for application users. The accelerated configuration also reduced the data center footprint and power and cooling requirements. With Dell servers, Dell storage, and virtualization, IT administrators can easily adapt to increasing performance demands from OLTP applications.

1 Hardware components

This section provides an overview of the hardware components used to test the workload.

1.1 Dell PowerEdge FX2 chassis

The PowerEdge FX architecture is a converged, modular, building-block concept that scales and adapts to varying workload demands. It is a scalable architecture that allows enterprises to grow their processing resources according to their business needs. The architecture allows the configuration to be workload optimized. Computing resources that are best suited for the workload can be selected to run data center applications.

The foundation of the FX architecture is the PowerEdge FX2 enclosure that combines compute, network, and storage in a single unit. Two configurations of the FX2 chassis are available: FX2 (unswitched) and FX2s (switched). The FX2 chassis does not have any PCIe slots while the FX2s chassis provides eight low-profile PCI Express (PCIe) Gen 3 expansion slots to extend connectivity to the servers.

The FX2 chassis can hold up to four, 1U, half-width servers. Server options include PowerEdge FC630 servers and PowerEdge FM120x4 microservers. Redundant, quad-port, 1GbE or 10GbE I/O modules are available on the chassis for server network connectivity. These I/O modules, along with three available FN I/O Aggregator modules, can simplify cabling, improve east/west traffic within the server, and enable LAN and SAN convergence. The chassis can be managed from each server using the agent-free iDRAC8 with Lifecycle Controller or on the chassis using the embedded Chassis Management Controller (CMC).

For more information, see the PowerEdge FX2 spec sheet.

1.2 Dell PowerEdge FC630 server

Designed for enterprises looking for high-performance computational density, the PowerEdge FC630 is a powerful workhorse for IT infrastructures. The half-width, two-socket server delivers exceptional performance in a compact form factor. The FX2 chassis can accommodate up to four FC630 servers, each supporting up to 18 cores with the latest Intel® Xeon® E5-2600 v3 processors and up to 24 dual in-line memory modules (DIMMs) for a maximum memory capacity of 768GB.

For more information, see the PowerEdge FC630 spec sheet.

1.3 Dell Networking FN410S IOA

The FN IO Aggregator is designed specifically for the PowerEdge FX2 converged-infrastructure chassis, part of the PowerEdge FX architecture. Supporting up to two FN IO Aggregators per chassis, the FX2 chassis also includes up to 8 x 10GbE internal ports, plus redundant cooling fans and power supplies.

The FN IO Aggregator can simplify FX2 connectivity by as much as 8-to-1, greatly reducing cabling complexity. Plug-and-play convenience with zero-touch and preconfigured features enables Level 2 control and allows fast, easy network deployment using a simple GUI, or optional management of the FN IO Aggregator through a CLI using the Dell Networking OS 9.

For more information, see the FN IO Aggregator page.

1.4 QLogic QLE2562 8Gbps host bus adapter

QLogic QLE2562 dual-port 8Gbps Fibre Channel (FC) to PCI Express host bus adapters (HBAs) are installed in the PCIe slots of the FX2 chassis. These HBAs provide connectivity between each of the FC630 servers and the storage array.

For more information, see the QLogic 2500 Series spec sheet.

1.5 Dell Storage SC4020 storage array

The Dell Storage SC4020 makes the business and technology benefits of true enterprise-class storage practical and affordable for mid-sized deployments. The SC4020 offers the award-winning capabilities of the popular Dell SC8000 array in a compact, affordable, yet highly expandable format. This easy-to-manage solution will auto-tune your datacenter environment to meet both workload and budget requirements at the same time.

Dual redundant controllers, 24 internal drives, eight ports of 8Gbps Fibre Channel network access, plus four IP ports for management and replication are all delivered in a space-saving 2U chassis, making the SC4020 a true all-in-one solution. The SC4020 supports 16GB of memory per module for a total of 32GB of memory per array. By adding Dell SC200 or SC220 enclosures, the SC4020 can support a raw capacity greater than 400TB.

With the Fluid Data architecture, Dell Storage changes the way organizations manage data. Empowered by real-time system information about each block of data, Dell Storage Center optimizes data placement, management and protection throughout its lifecycle. This level of data awareness also enables Storage Center to virtualize the storage infrastructure, bringing new efficiency, flexibility, and reliability to enterprise storage by creating a pool of high-performance storage shared by all servers and applications.

For more information, see the Dell Storage SC4020 page.

1.6 Dell Networking S4810 Ethernet switch

The Dell Networking S4810 switch is a 1U, rack-mountable, 10GbE switch with 48 ports. It features a 1.28Tbps (full-duplex), non-blocking switching fabric designed to deliver line-rate performance under full load with low application latency.

For more information, see the Dell Networking S-Series Managed Switches page.

1.7 Brocade 6505 Fibre Channel switch

The Brocade 6505 switch is a 1U, 24 port, rack-mountable Fibre Channel switch providing up to 16Gbps of bandwidth per port. This switch enables organizations to simplify IT infrastructures, improve system performance, maximize the value of virtual server deployments, and reduce overall storage costs.

For more information, see the Brocade 6505 Switch page.

2 Network configuration

Each PowerEdge FC630 server has an embedded dual-port Broadcom 10GbE network interface card (NIC). Using the Broadcom Advanced Control Suite, a network team is created on each server using the two embedded network ports. Through internal connections, each port in the FC630 is connected to each of the two FN410S IOA modules installed in the FX2 chassis. The FN410S IOA modules are connected to two S4810 top-of-rack (ToR) switches using 10GbE SPF+ networking cables. The FN410S modules and the S4810 switches were deployed in a multi-Virtual Link Trunking (mVLT) configuration to provide high availability for LAN traffic. The cabling diagram in Figure 1 shows how the networking components were connected to the FX2 chassis.

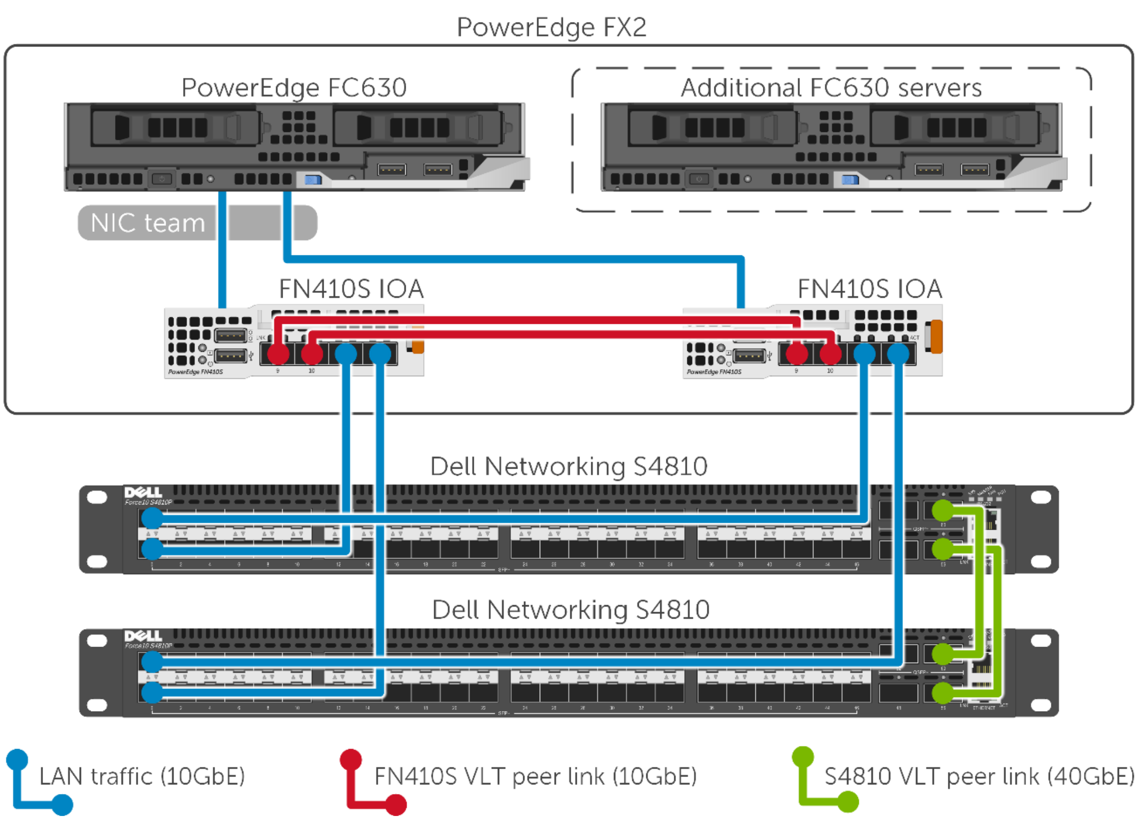

Figure 1 Network diagram

Instructions for deploying this configuration can be found in the section, "Environment Two: Dell Networking Switches with mVLT and IOA in Programmable MUX mode," in the document, [PowerEdge FX2 - I/O Aggregator - VLT Deployment Guide].

3 Storage configuration

The Fibre Channel version of the SC4020 was used to provide storage. Two external SC220 drive enclosures were added to the SC4020 in order to accommodate the drives used for testing. The internal drive bays and one external enclosure were populated with forty-eight 300GB 15K spinning drives. The second external enclosure was half-populated with twelve 1.6TB read-intensive SSDs.

3.1 Front-end configuration

The FX2 chassis used for testing included eight PCIe slots. Each server bay is connected to two of the PCIe slots. A total of four dual-port QLogic QLE2562 Fibre Channel HBAs, one for each FC630, were installed in the PCIe slots. The front-end ports on the SC4020 were configured to run in virtual port mode, in a dual-zone fabric, to provide a fault-tolerant connection to the servers. The cabling diagram in Figure 2 shows how the FX2 chassis was connected to the SC4020.

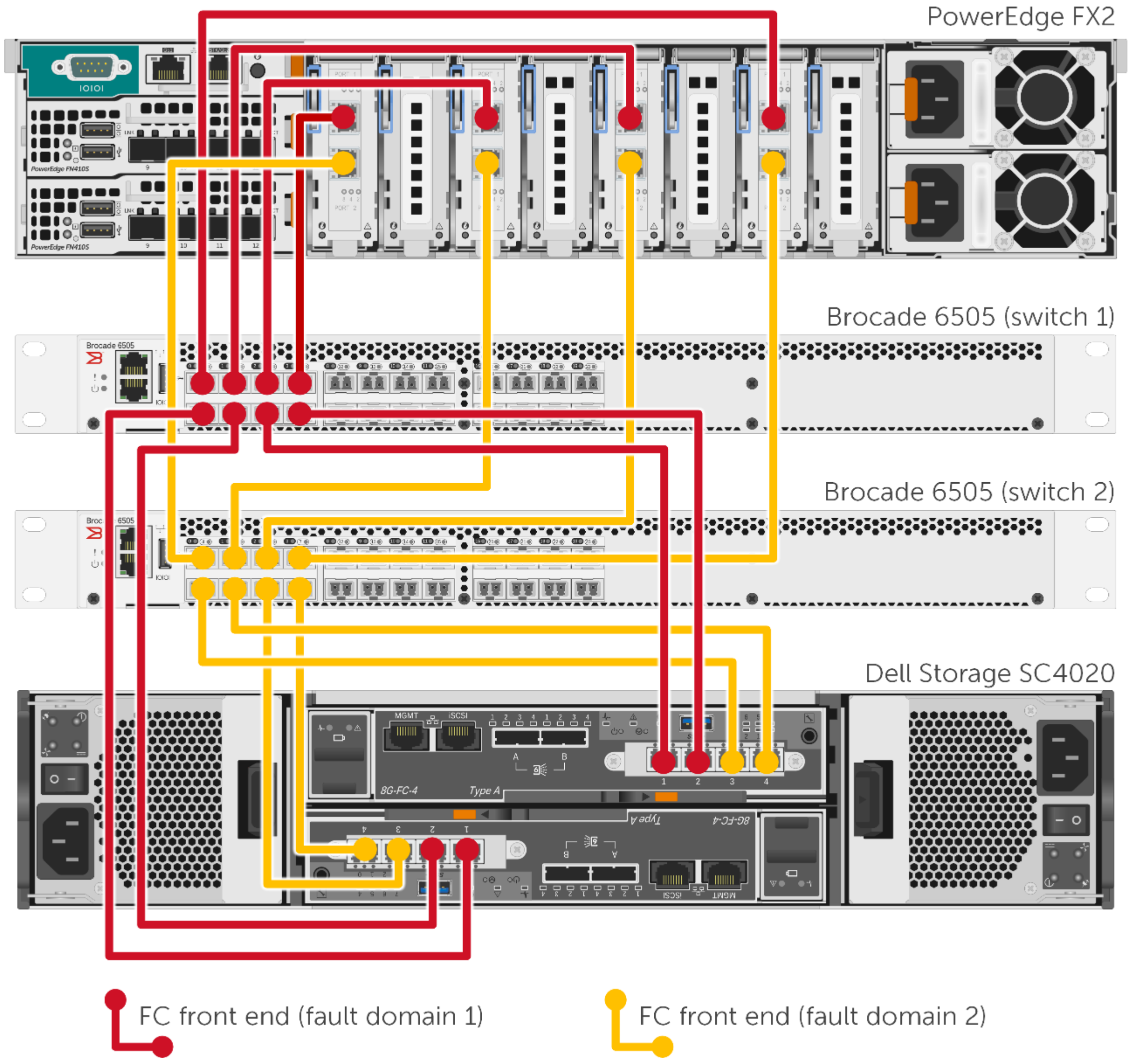

Figure 2 Front-end cabling diagram

3.2 Back-end configuration

A single SAS chain was used to connect the SC220 enclosures to the SC4020. The cabling diagram in Figure 3 shows how the SC220 enclosures were connected to the SC4020.

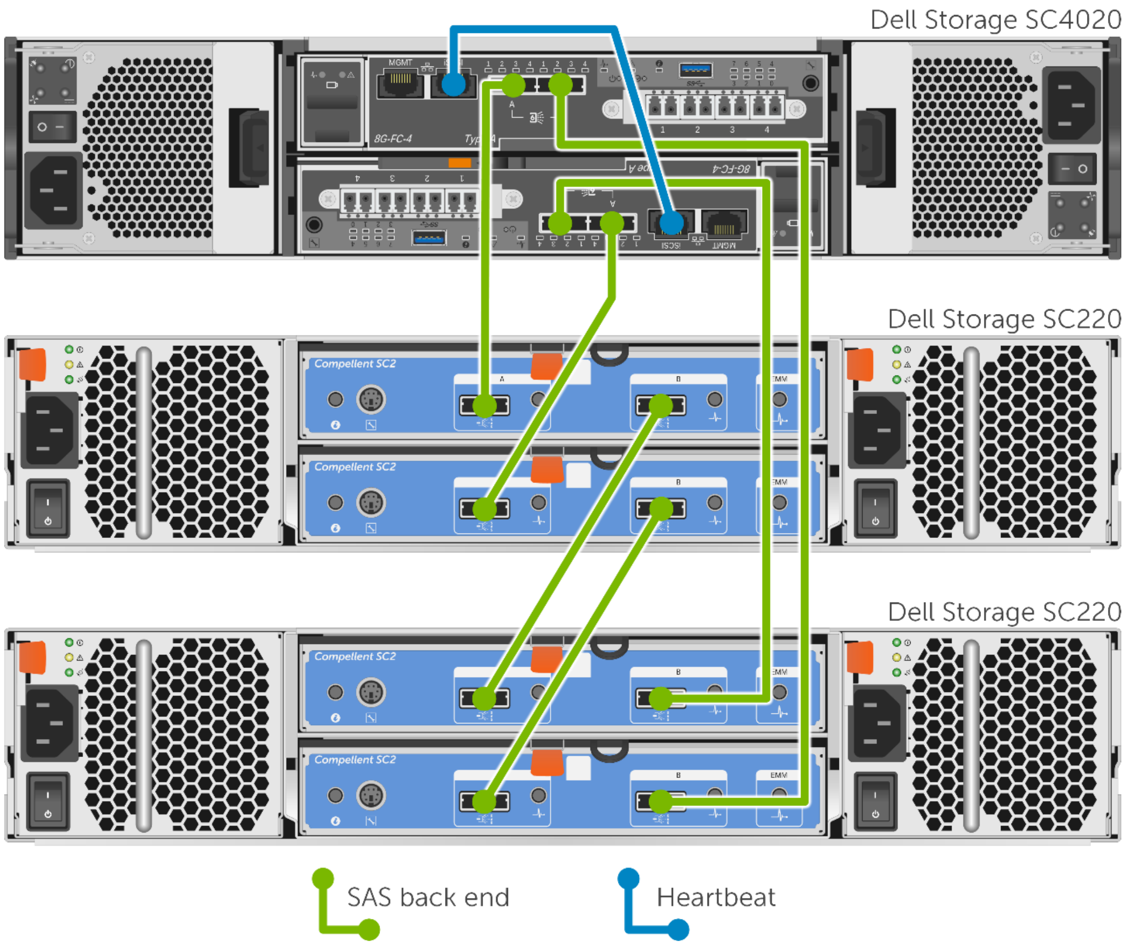

Figure 3 Back-end cabling diagram

4 Hyper-V cluster configuration

A Windows Server 2012 R2 Hyper-V cluster was created to host the virtual machines used for testing. Windows Server 2012 R2 was installed on each FC630 server, including the Hyper-V role and the Failover Clustering feature. A Windows failover cluster was then created using each server as a node in the failover cluster as shown in Figure 4.

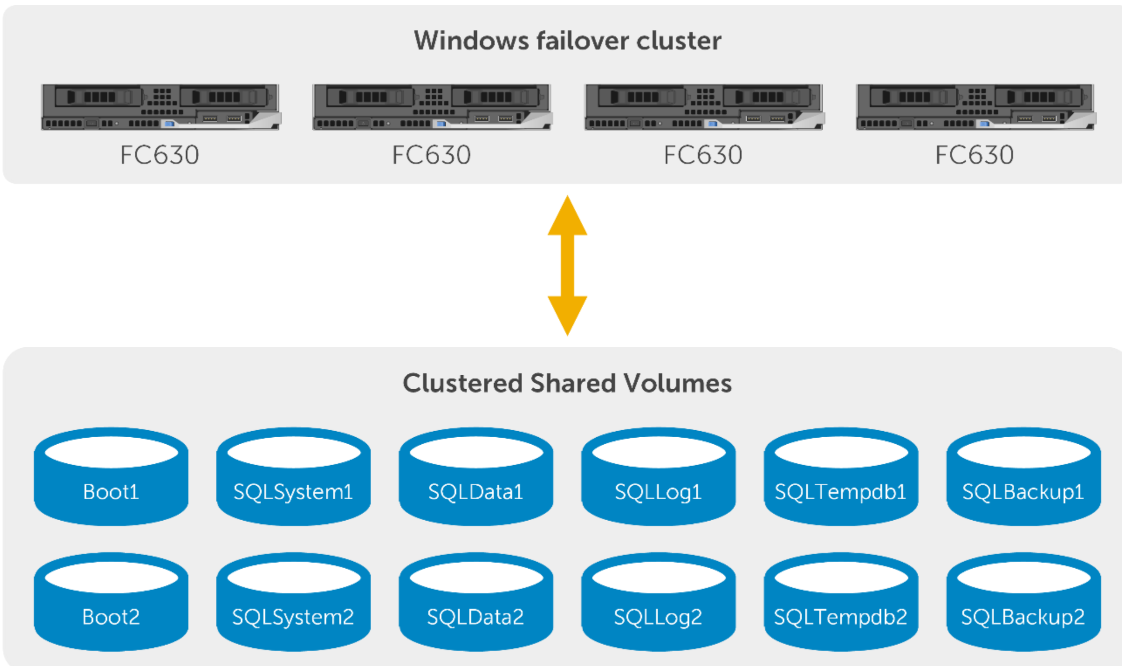

Figure 4 Hyper-V cluster configuration

Cluster configuration highlights:

- The system profile was set to **Performance** in the system BIOS on each server.
- Each server was configured to use the **High performance** power plan in Windows.
- The Multipath I/O (MPIO) feature was installed on each server to provide redundancy and optimal performance for each Storage Center volume used by Hyper-V. MPIO was configured using Dell Storage best practices as documented in the paper, *Dell Storage Center - Microsoft Multipath I/O Best Practices*.
- To provide storage for the application and database virtual machines, twelve 2TB Storage Center volumes were presented to the Hyper-V cluster, formatted with an allocation unit of 64KB and configured as Clustered Shared Volumes (CSVs).

- The CSVs were evenly distributed across the controllers in the Storage Center, as shown in Table 1.

Table 1 CSVs on each controller

CSVs on controller 1	CSVs on controller 2	CSV purpose
Boot1	Boot2	Windows boot volumes for VMs
SQLSystem1	SQLSystem2	SQL Server system databases for VMs
SQLData1	SQLData2	SQL Server database data files for VMs
SQLLog1	SQLLog2	SQL Server database log files for VMs
SQLTempdb1	SQLTempdb2	SQL Server tempdb data files for VMs
SQLBackup1	SQLBackup2	SQL Server backup files for VMs

- The Broadcom Advanced Control Suite was used to create a NIC team on the embedded NIC ports in each FC630. This provides higher bandwidth as well as fault tolerance for LAN traffic. Using Hyper-V Manager, a "vSwitch LAN" virtual switch was created on each node using the NIC team. All virtual machines used this virtual switch for LAN traffic.
- The CSVs were mounted in the C:\ClusterStorage folder on each node and renamed to reflect how they are used.

5 Virtual machine configuration

A total of eight virtual machines were created to generate the OLTP workload. Four applications servers and four database servers were created as generation-2 virtual machines. Each virtual machine was also configured for high availability, allowing the virtual machines to easily move between nodes in the Hyper-V cluster.

All virtual machines used VHDx files for storage. Disks that stored SQL Server database files were created as fixed-size disks. All other disks were created as dynamically expanding disks to conserve disk space.

Application server VM configuration highlights:

- Each VM was given 4GB of memory.
- Each VM was given 4 virtual processors.
- Each VM was given 1 virtual hard disk, which was presented as the C: drive.
- Windows Server 2012 R2 was installed.
- Benchmark Factory was installed to generate the OLTP workload.

Database server VM configuration highlights:

- Each VM was given 24GB of memory.
- Each VM was given 18 virtual processors.
- Each VM was given 7 virtual hard disks using the CSVs shown in Table 2.

Table 2 CSVs used by each database server VM

Drive letter	SQL1 CSV	SQL2 CSV	SQL3 CSV	SQL4 CSV
C:	Boot1	Boot2	Boot1	Boot2
D:	SQLSystem1	SQLSystem2	SQLSystem1	SQLSystem2
E:	SQLData1	SQLData1	SQLData1	SQLData1
F:	SQLData2	SQLData2	SQLData2	SQLData2
G:	SQLTempdb1	SQLTempdb2	SQLTempdb1	SQLTempdb2
H:	SQLBackup1	SQLBackup2	SQLBackup1	SQLBackup2
I:	SQLLog1	SQLLog2	SQLLog1	SQLLog2

- Windows Server 2012 R2 was installed with the following:
  - The **High performance** power plan was selected using the Power Options in Windows.
  - Using the Local Group Policy Editor, the following rights were granted to the SQL Server service account:
    - Lock pages in memory
    - Perform volume maintenance tasks

- SQL Server 2014 was installed with the following:
  - The following startup parameters were used:
    - **-T1117:** This flag forces autogrow to expand all files in a filegroup by the same amount.
    - **-T1118:** This flag forces tempdb to always allocate full extents.
  - The **max server memory (MB)** option was set to 21,500 MB.
  - The **max degree of parallelism** option was set to 1.
  - The tempdb database was configured to use eight 25GB data files and one 10GB transaction log file.

6 Test strategy

Performance metrics measuring application and storage performance were captured while executing the virtualized OLTP workload on the baseline hardware configuration. After baseline testing was complete, the hardware was upgraded to create the accelerated configuration. The same performance metrics were then captured while running the same workload on the new configuration. The effectiveness of the upgrades was measured by evaluating the performance of the workload on the two different hardware configurations.

6.1 OLTP workload

Instead of using a single application server and a single database server to generate the entire workload, four sets of application and database servers were used. This was done to simulate a virtual environment that has multiple smaller databases, each serving their own applications.

Benchmark Factory was used to generate the OLTP workload. A scale factor was chosen that would generate a database roughly 520GB in size. To help distribute the database load across the volumes on the Storage Center, the database was created with four 150GB data files. Since each database server has two volumes for SQL Server database files, two data files were placed on each data volume. The same database was used on each of the four database servers, providing a total of approximately 2TB of data for the workload.

The user load was set to 500 concurrent users on each application server. To increase the amount of work performed by each user, the think time and keying time for each transaction were reduced by a factor of 1000. This simulates an application that uses connection pooling, where application users share connections to the database, reducing the number of connections to the database server. Since there are four application servers, the workload simulates a total of 2000 concurrent users.

6.2 Baseline configuration

The baseline configuration was created as shown in Figure 5.

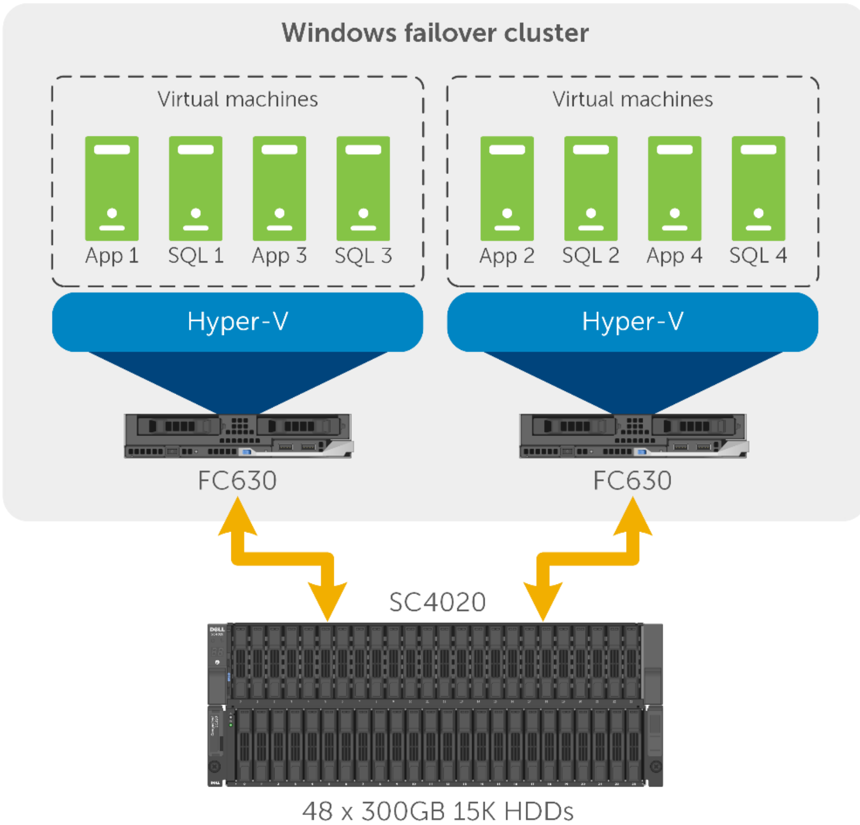

Figure 5 Baseline testing configuration

For the baseline testing, the Hyper-V cluster used two FC630 servers to host the virtualized application and database servers. To provide storage for the eight virtual machines, the SC4020 was configured to use a single tier of forty-eight 15K spinning drives. All volumes were provisioned from that tier. Each node in the Hyper-V cluster hosted two application servers and two database servers.

6.3 Accelerated configuration

After completing the baseline testing, the baseline configuration was modified to create the accelerated configuration shown in Figure 6.

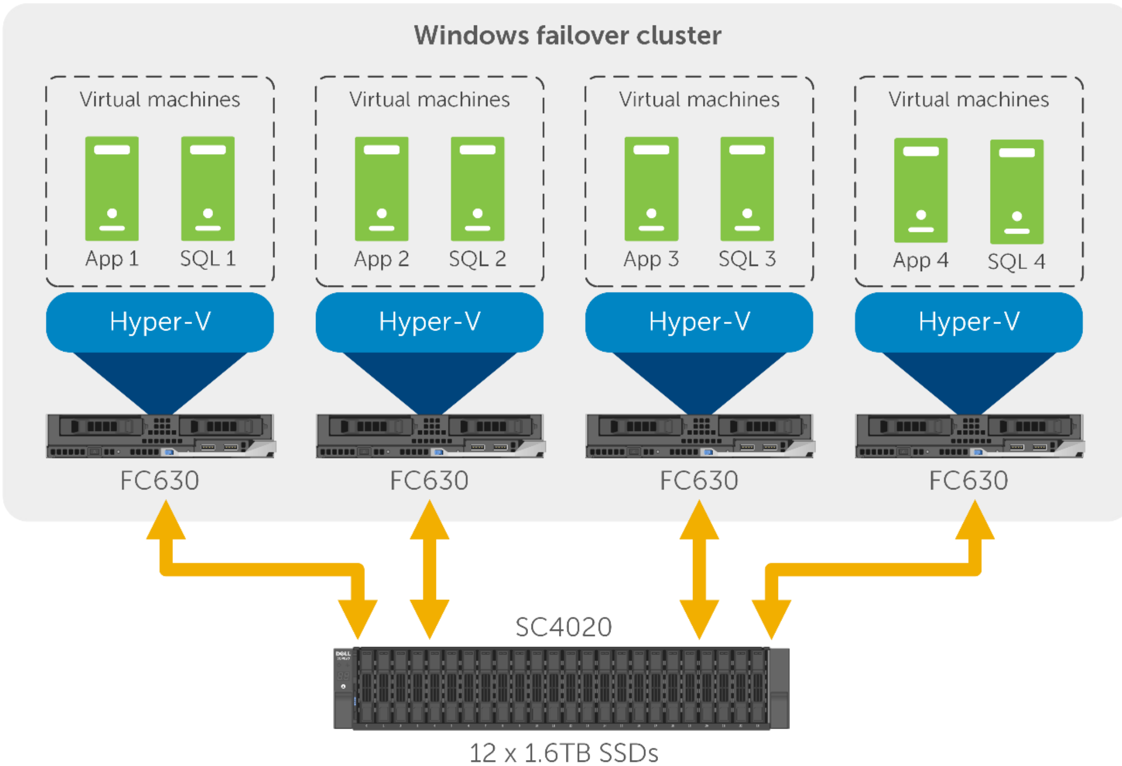

Figure 6 Accelerated testing configuration

To provide additional compute and memory resources, two additional FC630 servers were used by the Hyper-V cluster. The virtual machines were redistributed across the four cluster nodes so that each node hosted one application server and one database server. This allowed more resources to be allocated to the database servers. The number of processor cores was increased from 18 to 40 and the amount of memory was increased from 24GB to 52GB. No additional hardware resources were given to the application servers.

To improve storage performance, SSDs were used instead of spinning drives. All volumes were moved from spinning drives to a single tier of twelve 1.6TB read-intensive SSDs. No other storage configuration changes were made. Read-intensive SSDs were chosen instead of write-intensive SSDs due to their lower cost per GB. When compared to write-intensive drives, read-intensive drives offer comparable read performance but lower write performance. Because of the reduced write performance, the suitability of read-intensive drives for OLTP workloads will vary from environment to environment. The more write-intensive the workload is, the less suitable read-intensive SSDs will be.

The hardware upgrades required some SQL Server configuration changes. To use the extra memory given to the database servers, the **max server memory (MB)** setting was increased from 21500 to 48128.

Additional memory allows more data to be cached in SQL Server, which helps reduce the load on the storage subsystem. To help the read-intensive SSDs handle the increase in writes generated with the accelerated configuration, the **–k** startup parameter was used to limit the amount of data that checkpoints were allowed to write. The **–k50** startup parameter was added to each SQL Server instance to limit the writes generated by checkpoints to 50MB/s per instance.

7 Test results

To measure the success of the hardware changes made to the baseline configuration, key performance metrics were captured and compared. Performance metrics reported by each instance of Benchmark Factory were used to evaluate application performance. Storage performance was evaluated using metrics captured by Windows Performance Monitor (PerfMon) on each of the database servers. Metrics were aggregated across the virtual machines generating the workload to measure the entire workload.

7.1 Application performance

Poor application performance is not only frustrating to end users, it can cost businesses money. Slow applications not only reduce the effectiveness of employees, it can also affect revenue when those applications limit the ability to sell products to customers.

By running the workload on the accelerated configuration, there was a significant improvement in application performance. The applications performed more transactions with a much lower response time.

7.1.1 Transactions per second

To measure how much work is being performed by the application, the number of transactions per second (TPS) was captured. After upgrading the baseline configuration to the accelerated configuration, the number of transactions per second increased by 650 percent.

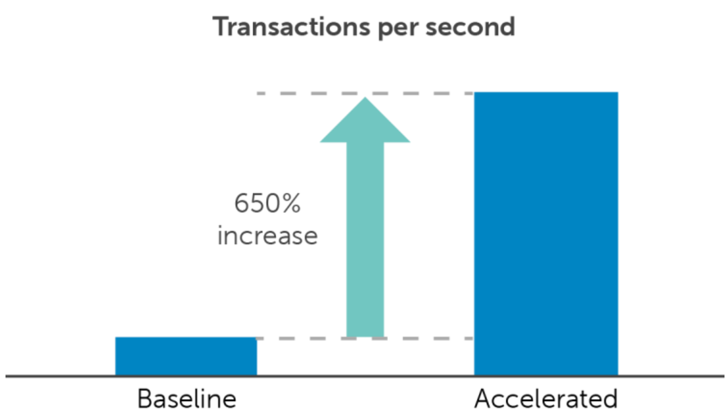

Figure 7 Transactions per second compared to baseline configuration

7.1.2 Average response time

A lower application response time improves the end user's experience and may be noticed much more quickly than an increase in transactions per second.

After applying hardware upgrades, the decrease in application response time was displayed in dramatic fashion with a 95 percent reduction.

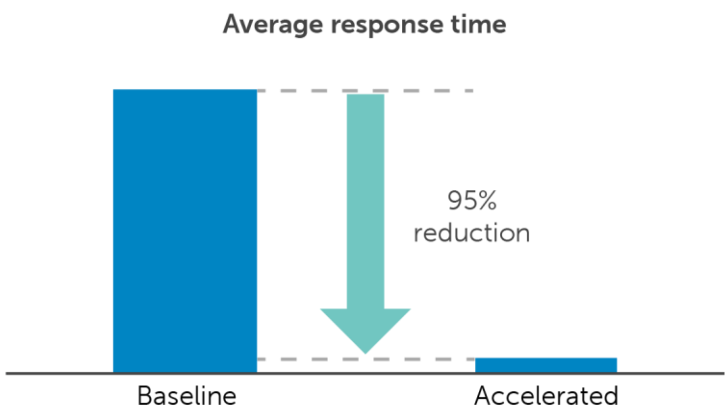

Figure 8 Average response time compared to baseline configuration

7.2 Storage performance

OLTP applications are characterized by having frequent, small transactions that randomly access the database. Because OLTP applications typically do not perform well without good storage performance, high-performance storage is critical for databases servicing this type of workload.

By running the workload on the accelerated configuration, there was a significant performance improvement and a reduction in drive count by 75 percent, allowing for lower costs with a reduced datacenter footprint and lower power and cooling requirements.

7.2.1 IOPS

When testing the baseline configuration, the database was limited by the performance of the I/O subsystem. This was easy to see by looking at the wait statistics inside of SQL Server. Solid state disks (SSDs) can provide lower latency at much higher I/O loads than spinning drives. By moving the database to

SSDs, the database was able to perform many more I/O operations — more than twice the number of IOPS on the accelerated configuration.

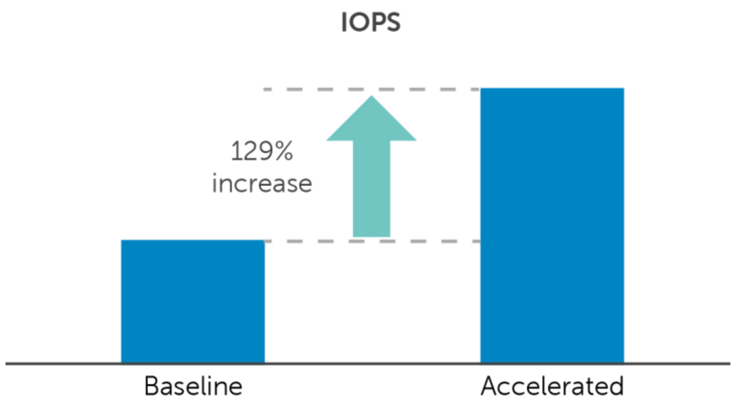

Figure 9 IOPS compared to baseline configuration

7.2.2 Average latency

A good measure of storage performance is the amount of time it takes to service an I/O request, which is also known as disk latency. High disk latency is detrimental to database performance. By using SSDs in the accelerated configuration, latency dropped by 99 percent.

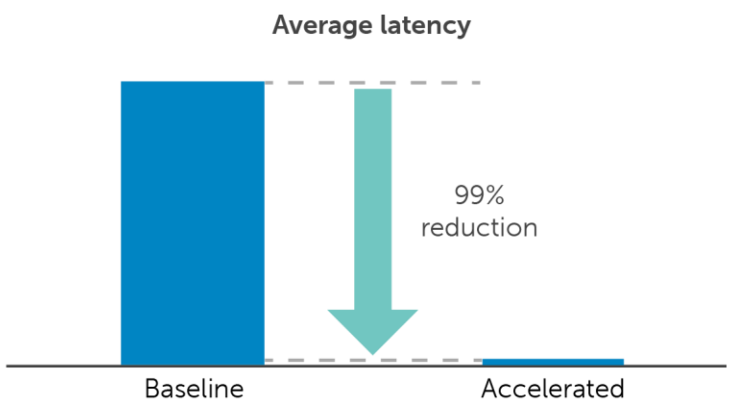

Figure 10 Average latency compared to baseline configuration

8 Conclusion

The performance of OLTP applications depends on having sufficient hardware resources. When virtualized on an FX2 chassis, CPU and memory resources can be increased by simply adding additional servers to the FX2 chassis. The SC4020 storage array makes it easy to increase and boost storage performance by replacing spinning drives with SSDs. With Dell servers and Dell storage arrays, you can easily accelerate your OLTP applications to satisfy current and future performance requirements while simultaneously lowering your data center footprint as well as reducing power and cooling requirements.

A Test component specifications

A.1 Hardware specifications

Hyper-V hosts	
Model	PowerEdge FC630
CPU	2 x Intel Xeon E5-2690 v3 @ 2.6Ghz
Memory	8 x 8GB DDR4 2133 MT/s (64GB total)
Network	Integrated Broadcom 57820s dual-port 10GbE
Host bus adapter (HBA)	QLE2562 dual-port 8Gb FC

Fibre Channel switch	
Model	Brocade 6505
Ports	24
SFPs	8Gbps

Storage array	
Model	SC4020
Protocol	Fibre Channel
External enclosures	1 x SC220

A.2 Firmware versions

Hardware	Component	Version
PowerEdge FX2	CMC	1.10.200.201410300014
PowerEdge FC630	BIOS	1.0.3
	Backplane	2.20
	Broadcom BCM57810	7.10.18
	Dell 64 Bit uEFI Diagnostics	4239A17

Hardware	Component	Version
	Dell OS Driver Pack for 13G	14.11.01
	Disks (internal)	LS0A
	Integrated Dell Remote Access Controller (iDRAC)	2.05.05.05
	Lifecycle Controller	2.05.05.05
	PERC H730 Mini	25.2.1.0037
	QLogic QLE2562	03.21.04
	System CPLD	0.3.6
Dell Storage SC4020	SCOS	6.5.20
Dell FN410S	FTOS	9.6(0.0)
Dell Networking S4810	FTOS	9.6(0.0)
Brocade 6505	Firmware	7.2.0a

A.3 Software versions

Hyper-V host	Version
Microsoft Windows Server 2012 R2 Standard Edition	6.3.9600
Broadcom Drivers and Management Applications	16.6.7.2
Broadcom Advanced Control Suite	16.4.3.0

Application server (Hyper-V VM)	Version
Microsoft Windows Server 2012 R2 Standard Edition	6.3.9600
Dell Software Benchmark Factory	7.1.1.630

Database server (Hyper-V VM)	Version
Microsoft Windows Server 2012 R2 Standard Edition	6.3.9600
Microsoft SQL Server 2014 Enterprise Edition	12.0.2000

B Additional resources

B.1 Technical support and resources

For Copilot support of Dell SC Series products:

- Dell SC Series customer portal (requires account)
- Email: support@compellent.com (non-emergency business hours)
- Phone: 866-EZ-STORE (866-397-8673) (United States only)
- Dell global online support

The Dell SC Series customer portal is an online portal for existing customers. A valid portal account is required to access the Knowledge Center. Once login to the portal, go to “Knowledge center”.

Dell TechCenter is an online technical community for IT professionals and is a great resource to discover and learn about a wide range of technologies such as storage, servers, networking, software, and cloud management.

B.2 Related documentation

Table 3 Referenced or recommended resources

Vendor	Resource
Dell	PowerEdge FX2 spec sheet
Dell	PowerEdge FC630 spec sheet
Dell	FN IO Aggregator page
Dell	Dell Storage SC4020 page
Dell	Dell Networking S-Series Managed Switches page
Dell	PowerEdge FX2 - I/O Aggregator - VLT Deployment Guide
Dell	Dell Storage Center - Microsoft Multipath I/O Best Practices
Dell	Dell SQL Server Solutions
Dell	Dell Networking technical content on Dell TechCenter
Dell	Dell Storage technical content on Dell TechCenter
Brocade	Brocade 6505 Switch page
QLogic	QLogic Fibre Channel Adapters page