# Why Dell EMC Storage for Microsoft SQL Server

Eliminate inefficient data silos while increasing performance and agility for optimal performance and faster insights.

Microsoft SQL Server 2019: New and Enhanced Features

IDC predicts that by 2025, more than 163 zettabytes of data will be created in that one year alone challenging companies to not only effectively manage data on this massive scale but leverage it through analytics to turn it into data capital. Managing data at this scale has strong implications for IT infrastructure. Requirements around performance, capacity expansion, availability, and automation are all essential.

When we look historically at Microsoft SQL Server workloads we see the power of a structured data environment. Going forward, Microsoft continues to innovate and shape the role of SQL Server in the modern datacenter. With the introduction of big data clusters in Microsoft SQL Server 2019, the steps to gain insights on the structured and unstructured data within an environment have been greatly improved. Having the ability to gain insights on data wherever it sits is a game changing capability, however, this does not mean that data is best accessed wherever it may sit today.

Starting with SQL Server 2016, SQL Server has been expanding past structured data to also draw insights from unstructured data with an increased number of connectors to data sources. These data virtualization, PolyBase features allow for performance improvements by analyzing the data where it sits and then only sending relevant data back to SQL Server.
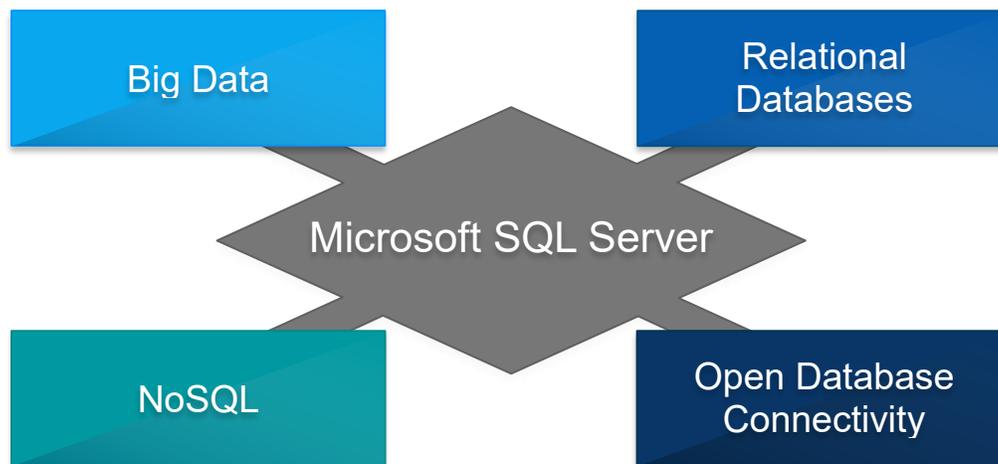
SQL Server 2019 includes a broad range of technologies, features, and services, supporting mission-critical applications such as analytics, in-memory databases, business intelligence (BI), and reliable and scalable online transaction processing (OLTP). The SQL Server platform has acquired capabilities to handle data integration, data warehousing, reporting, high-speed advanced analytics, data replication and programmability features including hosting in-database common language runtimes, service broker hosting, and semi-structured datatype management.

SQL Server database environments continue to grow in both size and complexity. With SQL Server 2019, Microsoft improved SQL Server core features and added new ones such as support for Big Data workloads with Apache Spark and HDFS. Dell EMC, in partnership with Microsoft, continues to provide the essential infrastructure components—servers, storage, and networking—for your SQL Server environment.

The goal of this paper isn't to provide an exhaustive list of features of Microsoft SQL Server 2019 or its benefits, but rather how Dell EMC Storage is providing the optimal storage portfolio to enable DBAs, application owners and IT decision makers to maximize their investments in Microsoft SQL Server.

The right data, in the right place, at the right SLA

IDC predicts that by 2025, more than 163 zettabytes of data will be created in that one year alone challenging companies to not only effectively manage data on this massive scale but leverage it through analytics to turn it into data capital. Managing data at this scale has strong implications for IT infrastructure. Requirements around performance, capacity expansion, availability, and automation are all essential.

When we look historically at Microsoft SQL Server workloads we see the power of an efficiently structured data environment. Going forward, Microsoft continues to innovate and shape the role of SQL Server in the modern datacenter. With the introduction of big data clusters in Microsoft SQL Server 2019, the steps to gain insights on the structured and unstructured data within an environment have been greatly improved. Having the ability to gain insights on data wherever it sits is a game changing capability, however, this does not mean that data is best accessed wherever it may sit today.

For the past few years, the migration to flash storage from spinning disk has been a major focus in the enterprise data center as exabytes (EB) of data have been migrated from spinning disk to flash storage. The eventuality of this migration should not have been surprising as the potential performance gains offered by flash storage are orders of magnitude greater than that of spinning disk.

## Consolidation for Optimized Microsoft SQL Server Environments
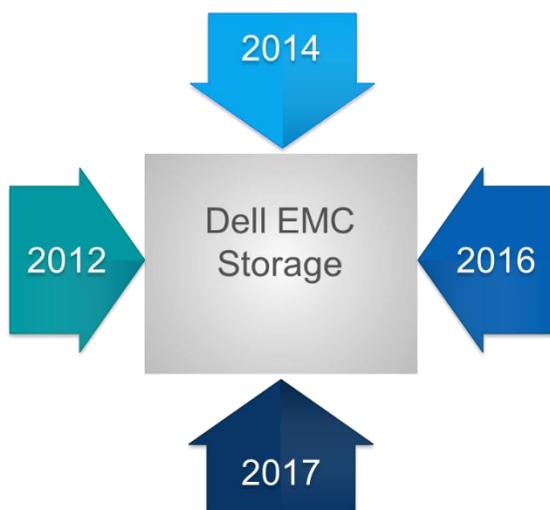
Dell Technologies is leading the modernization of the datacenter with Dell EMC storage solutions that allow for the consolidation of data on platforms designed for 99.9999% of availability, sub-millisecond sustained latency, maximum agility, and superior security. Optimizing the underlying infrastructure for Microsoft SQL Server not only benefits the storage administrators; it also has a significant impact on the effective agility of DBAs leveraging these arrays.

The reality for most companies is that their Microsoft SQL Server environments span many different versions which have disparate features and capabilities. Consolidation of these versions on to a modern active-active all-flash array allows for the offloading of responsibilities to the infrastructure and helps to provide a more consistent and elevated experience to all versions. The benefits that are experienced are largely based on the data services provided by the array and these are highly differentiated across the industry with respect to the performance impact and effectiveness of the data services.

Modern Array Data Services

In an all-flash environment there is always significant focus on the effective capacity of the array to maximize efficiency and TCO. While creating these efficiencies of how data is stored on an array, a side effect is unparallel agility for your SQL Server environment. Across the Dell EMC storage product portfolio there are also several key differentiators and variations on features; this section will focus on the commonalities of Dell EMC primary storage arrays.

For the purpose of this paper, those include Dell EMC Unity XT, PowerMax, XtremIO X2 and SC.
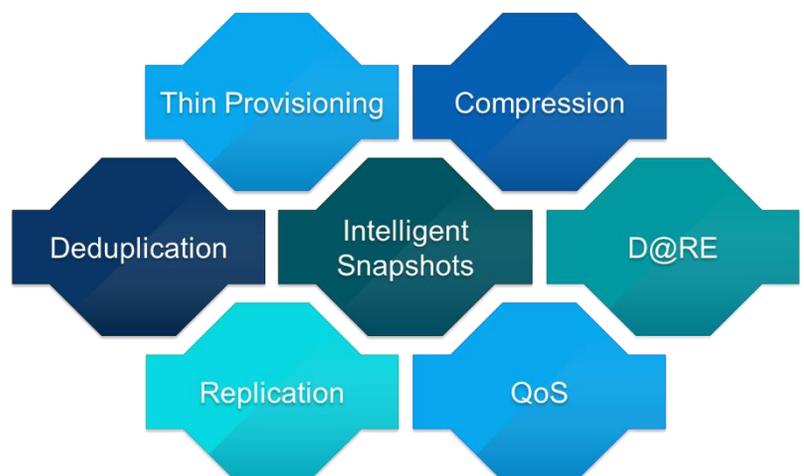
Usable Capacity: An often-overlooked aspect is the impact of system overhead, or lack thereof, that the management, monitoring and the underlying RAID create. Dell EMC arrays provide some of the highest usable capacity in the industry with systems providing as much as 88% usable capacity to the servers.

Thin Provisioning: Commonplace today, this is one of the key features for maximizing the cost effectiveness of a shared storage environment. For a DBA, this allows them to allocate the database files to the expected size in the future and not have to increase the size of it later. Space is only consumed as the data is written to the array

Compression: While this is done in slightly different ways on the Dell EMC portfolio of storage arrays, each system typically will provide better than 2:1 compression on a database that is not already leveraging row or page compression within Microsoft SQL Server. Relying solely on the storage array to provide compression often can result in the improvement of performance due to the freeing up of SQL Server CPU resources. Testing should be done in all cases, however, since the impact can vary depending on workload.

Deduplication: There are many reasons a DBA or application owner would need to create database copies. Test/Dev, patch testing, near real-time analytics, or just making a copy for protection, these database copies will be 100% deduplicated on the array. Only the delta unique changes that are made to the copies or production will result in new data being written to the array.

Intelligent Snapshots: As efficient as deduplication is, an even more efficient and faster way to make a database copy is to simply take a snapshot of a volume or set of volumes and mount the snapshots to a SQL Server instance. Making database copies in this way ensures little to no impact to the production database and allows copies of even very large (20TB+) databases to be provided in under 5 minutes.

D@RE: Data at Rest Encryption ensures that all sensitive user data stored on array is encrypted when being written to disk, so that private data does not fall into the wrong hands. With simple activation, management, and auditing functionality, D@RE is a powerful tool to protect data, regardless of an organization's size, industry, or use case.

Replication: Array-based replication of Microsoft SQL Server databases can easily co-exist with database replicas created by AlwaysOn Availability Groups. This replication offloads burden from SQL Server to manage the replication without impact to the host. It also allows you a single method of replication that is consistent regardless of the Microsoft SQL version.

QoS: Quality of Service, even in the world of all flash arrays, can play an important role to easing the concerns of some organizations to take advantage of consolidation of production, test/dev, analytics and more.

iCDM: Integrated Copy Data Management for Microsoft SQL Server

With all those great data services as a foundation, we can now focus on one of the biggest issues facing a database environment. Copies of production databases are required by many different stakeholders and they want them for different reasons:

- Developers use copies as source for the next iteration of the application, to help identify and fix bugs in current applications.
- All applications need to be backed up at some point, so data protection copies are required.
- Reporting functions need copies to enable better business decisions.

Copies enable businesses to allow these functions to run in parallel without interfering with production processing.

Self
Service

Automation &
Orchestration

Intelligent Snapshots

Dell EMC Storage
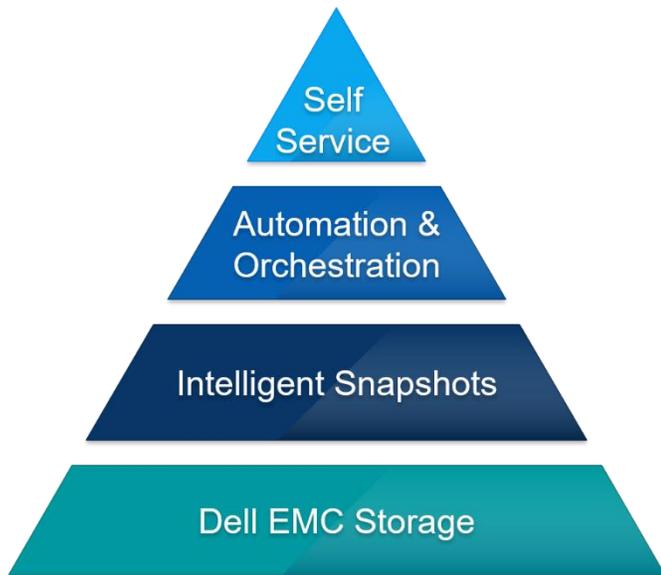
iCDM addresses copy management use cases in a consolidated SQL database environment for database protection, recovery and database repurposing. Dell EMC AppSync™ automatically discovers application databases, learns the database structure, and maps it through the virtualization layer to the underlying Dell EMC storage.

With copies of databases being provisioned so easily via snapshots, consolidation of production and test/dev environments on the same all flash array can provide significant benefits. Dell EMC AppSync simplifies, orchestrates, and automates the process of generating and consuming application consistent copies of production data.

AppSync's deep application integration coupled with abstraction of underlying Dell EMC storage and replication technologies empowers application owners to satisfy copy demands for data repurposing, operational recovery and disaster recovery, all from a single user interface.

- Simple: Automated copy management without the need for custom scripts.

- Intelligent: Tight integration with applications, hosts, virtualization, Dell EMC storage and replication environments.

- Frictionless: DBAs, applications owners and storage admins are on the same page with a transparent copy workflow

AppSync orchestrates all the activities required, from copy creation and validation through mounting the snapshots at the target host and launching or recovering the application. This solution supports and simplifies Microsoft SQL Server workflows that include refreshing and expiring copies as well as restoring the production database.

In addition to the orchestration options, iCDM allows Test/Dev engineers to be able to refresh their environment from the latest gold master copy of the database in an ad-hoc fashion whenever they need to do so. This results in accelerated development and optimization of the data reduction rate of the array since only the unique data blocks that are deltas between production, gold master, and the additional copies will result in storage growth.

How Do Containers Change the Landscape?

To improve competitiveness in a fast-paced world, organizations are embracing software as a prime differentiator. A software-driven business is more agile and efficient, innovates faster and responds dynamically to changing market and customer demands. Increasingly, gaining these compelling advantages requires adopting cloud-native applications. The benefits of containerized Microsoft SQL Server lie in the efficient provisioning, startup times, DevOps improvements, and being able to significantly reduce the overall storage capacity versus a VM.

Kubernetes is the de facto standard for container orchestration for microservices and applications. However, enterprise adoption of big data and databases using containers and Kubernetes is hindered by multiple challenges, such as complexity of persistent storage, availability, and application life-cycle management. Kubernetes provides the agility and scale that modern enterprises require. However, Kubernetes provides the building blocks for infrastructure, but not a turnkey solution.

Originally containers and Kubernetes were optimized for stateless applications, restricting enterprises to use container technology for workloads, such as websites and user interfaces. More recently, stateful workloads have been increasingly run on Kubernetes. Whereas it is relatively easy to run stateless microservices using container technology, stateful applications require slightly different treatment.
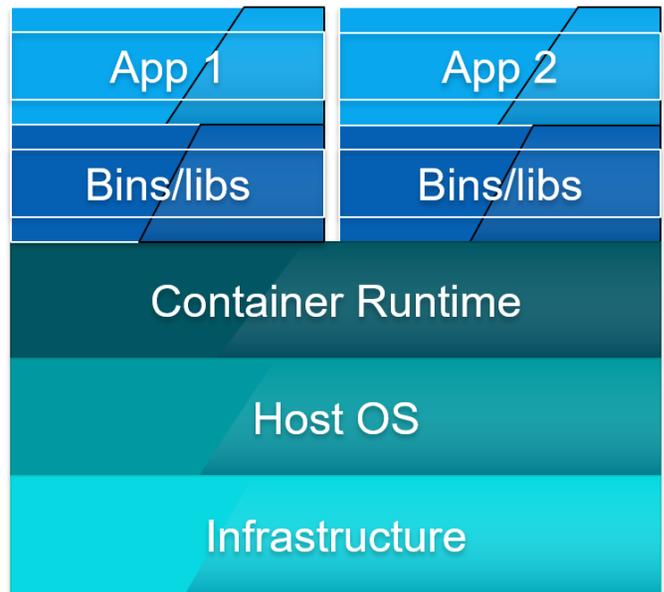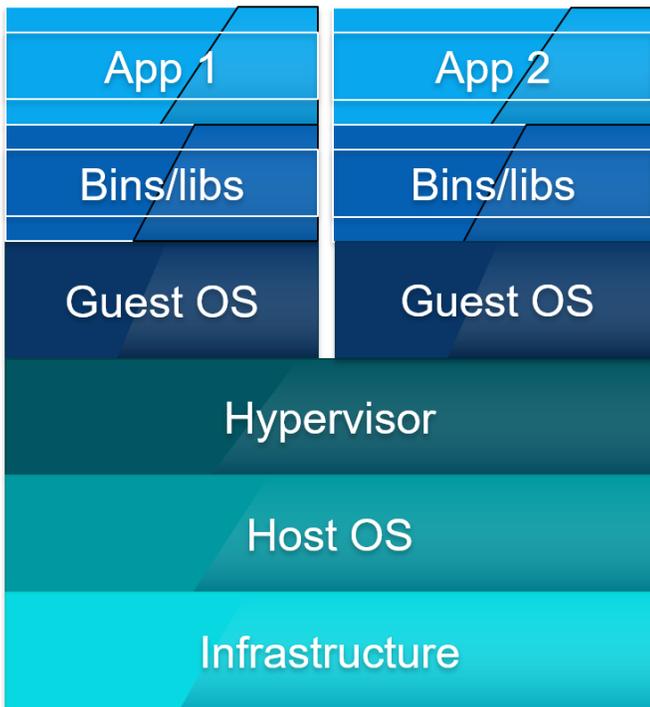
Container technology enables Microsoft SQL Server development teams to quickly provision isolated applications without the traditional complexities. For many companies, the use of containers starts with the departments that are focused on software development, to boost productivity and time to value. The journey typically starts with installing, implementing, and using containers for applications that are based on the microservice architecture.

Whether customers are looking to use containers for test and development purposes, like DevOps, or to deploy databases with container orchestrators such as Kubernetes, SQL Server in containers ensures a consistent, isolated and reliable behavior across environments, that is easy to use, start and stop. Customized content can be built on top of SQL Server containers, and run without being affected by the rest of the environment. This isolation makes SQL Server in containers ideal for test deployment scenarios as well as DevOps processes.

Dell EMC CSI Support

Containers are ephemeral by nature, so the data that needs to be persistent must survive through the restart/re-scheduling of a container. When containers are rescheduled, they can die on one host, and might get scheduled on a different host. In such cases, the storage should also be shifted and made available on a new host for the container to start gracefully. This is functionality is delivered via the CSI plugin.

Dell EMC brings the power of persistent storage to a stateless platform. Dell EMC's storage and CSI plug-ins enable container orchestrators, like Kubernetes, to easily provision highly available and scalable container volumes for stateful containerized applications and databases. CSI driver capabilities are constantly being improved upon so, be sure to leverage the latest version and follow along with updates on GitHub.com/Dell.

How Big Data Clusters change the game
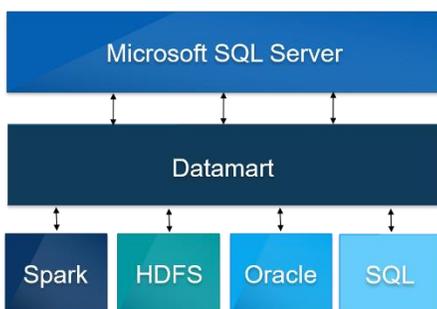
One of the most exciting new capabilities with the release of Microsoft SQL Server 2019 is Big Data Clusters. Big Data Clusters leverage containers and allow you to combine high-value data and high-volume data together into a single platform powered by SQL Server and other industry leading technologies such as HDFS and Spark. With multiple data formats, multiple sources, and increased data capacity and analytics capabilities, storage is more critical than ever.

Big Data Clusters allow existing SQL Server skillsets, like knowledge of T-SQL, to be used to query data that may exist in structured data sources such as Oracle and

SQL Server together with unstructured data. Silos of data that exist in the datacenter can be taken advantage of now via Azure Data Studio whether data sits on-premises or in the cloud.

The impact that this will have for businesses seeking to gain faster insights on their data is enormous. While some reporting processes still may require intensive ETL process, this allows for a fine tuning of data access through the sharded implementation where limited data is transferred, and insights are drawn closer to the source of the data.

Protecting Big Data Clusters

Big Data solutions are relatively new and while they are many different combinations and deployments, most face a common issue. The amount of data stored in these environments is in the terabytes to petabytes, but backup and recovery is often an afterthought. The size of the data combined with multiple sources can make them a real challenge to backup and recover. Common data protection strategies include rebuilding and rehydrating the big data environment, or maintaining parallel copies, both of which can be costly and time-consuming. As these Big Data Solutions become increasingly critical to business, the availability requirements will inevitably increase making these strategies unacceptable.

The storage snapshot capabilities of Dell EMC storage allow these environments to be protected by storage snapshots in a matter of seconds, allowing for fast, efficient point-in-time copies, and recovery times in minutes, not days or months.

No Compromise Mid-Range Storage for Microsoft SQL Server

Dell EMC Unity XT arrays combine compelling simplicity, all-inclusive software, blazing speed, optimized efficiency, and multi-cloud enablement to meet the needs of resource-constrained IT professionals in large or small companies. Designed for performance, optimized for efficiency and built for hybrid cloud environments, these systems are the perfect fit for supporting demanding virtualized applications, deploying unified storage and consolidating SQL Server environments.

Simplify the Data Landscape

Microsoft SQL Server's storage can be rightsized for performance, capacity and price at initial deployment and scale to meet future needs. Inline data efficiencies lower TCO while integrated copy data management provides orchestration and self-service for DBAs. Dell EMC's Unity XT is a scale-up architecture, available in variety of models to meet the needs for Microsoft SQL Server database deployments

Dell EMC Unity XT is as simple and easy to install as a consumer networked printer, increasing productivity. Simplify database storage management with an intuitive interface and leading virtualization integrations to support Microsoft SQL deployments. It only takes 25 minutes from opening the box until the Dell EMC Unity XT array is configured and ready to deliver operational productivity. As databases grow, capacity expansion is simple, scalable in increments as small as just 2 drives. Dell EMC Unity XT systems are 85% efficient and deliver up to 790TB effective capacity per rack unit.

Today's reporting processes require actionable insight to be drawn from a variety of data sources. Dell EMC Unity XT is both a block and file platform, allowing for the consolidation of the valuable data required for modern reporting with SQL Server Analysis Services.

Optimized Performance & Data Services

Dell EMC Unity XT is the simple, modern, flexible, and affordable storage solution for IT organizations looking to consolidate their Microsoft SQL Server database applications. Dell EMC Unity XT's architecture supports scores of VMs while delivering abundant IOPS at sustained low latency (<1ms) for running demanding Microsoft SQL workloads.

Dell EMC Unity XT combines inline compression and inline deduplication which combine to provide up to 5:1 data reduction with high resource efficiency that lowers TCO. Leveraging compression on the array instead of row or page compression at the database level can help improve CPU utilization and improve performance.

Rest Easy with Highly Certified Storage

All DBAs are diligently concerned with the overall security of their data, in-flight and at rest. Dell EMC Unity XT eases this concern, as it is Dell EMC's most federally certified storage and is on the US DoD approved products list. Dell EMC Unity XT meets database security requirements with D@RE, TLS, and IPv6 dual stack certification and more.

Enterprise Scale with Dell EMC PowerMax

The true performance potential of flash storage in the enterprise data center has been somewhat hamstrung by choke points around the input/output (I/O) path from the application to the storage. Fortunately, choke points lead to innovations. The Dell EMC™ PowerMax family is the first Dell EMC hardware platform that uses an end-to-end NVMe architecture for customer data. NVMe combined with higher-performing Storage Class Memory (SCM) media technologies delivers unmatched levels of performance for high value, high demand Microsoft SQL Server workloads.
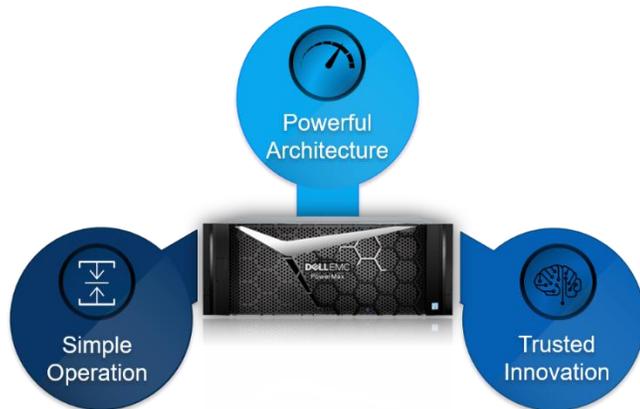
The innovative PowerMax reaches unprecedented I/O densities and performance by eliminating the flash media choke points found using traditional SAS and SATA interfaces. The PowerMax opens the door for Dell EMC customers to deploy innovative Microsoft SQL Server based applications in the areas of real-time analytics, machine learning, and big data that demand lower latency and higher performance.

For the most demanding Microsoft SQL Server environments, performance requirements may require a storage array to scale-out to provide enough resources to ensure the application is operating as efficiently as possible. A built-in machine learning engine (ML/AI), with automated data placement across SCM and flash drives, optimizes storage performance using pattern recognition and automated data placement without overhead.

Reliability, Availability, and Serviceability (RAS)

PowerMax's industry leading reliability, availability, and serviceability (RAS) make it the ideal platform for environments requiring always-on availability. These arrays are designed for six-nines of availability in the most demanding, mission-critical environments. All components are fully redundant to withstand any component failure, meaning there is no single point of failure.

The complete list of RAS features is extensive, with many features, inclusive of continuous system monitoring and advanced diagnostics to ensure the system provides optimal availability. A key feature available on PowerMax is SRDF (Symmetrix Remote Data Facility), the gold standard for remote replication. 70% of Fortune 500 companies use this tool to replicate their critical data to geographically dispersed data centers throughout the world. SRDF offers customers the ability to replicate tens of thousands of volumes, with each volume being replicated to a maximum of four different locations globally.

Reduce Backup Windows

Dell EMC PowerProtect for PowerMax dramatically reduces the database backup window by completely eliminating backup servers and the requirement to send backups over the network. Microsoft SQL Server DBAs can back up databases from SQL Server Management Studio with data sent directly from PowerMax to a Dell EMC Data Domain protection storage system. PowerProtect eliminates backup impact on application and database servers to gain up to 20 times faster backup, up to 10 times faster recovery, with reduced cost and complexity.

Go Beyond Quality of Service

Service levels for PowerMax ensure that applications have consistent and predictable performance by allowing users to separate storage groups based on performance requirements and business importance. Set specific service levels to ensure the highest-priority application response times are not impacted by lower-priority applications. The available service levels can be applied at the creation of a storage group or can be modified to an existing storage group at any time.

The capacity and performance of PowerMax arrays provide a single-tier system designed to provide high availability to keep up with increasing demand of data access. Having a single tier of flash storage, users need prioritized data access and have the ability to prioritize critical, high-priority applications while managing lower-priority applications. PowerMax

provides this ability with the service levels feature to allow users to set expectations for applications in order to provide predictable and consistent performance.

Isilon

Unstructured data sources are incredibly valuable sources that can be virtualized and then queried like any other table in SQL Server. In a Hadoop implementation on an Isilon cluster, Isilon OneFS serves as the file system for Hadoop compute clients. The Hadoop distributed file system (HDFS) is supported as a protocol, which is used by Hadoop compute clients to access data on the HDFS storage layer.
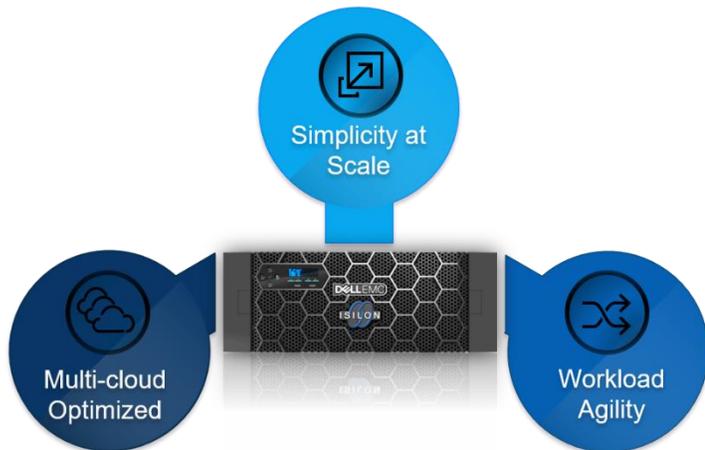
Hadoop compute clients can access the data that is stored on an Isilon cluster by connecting to any node over the HDFS protocol, and all nodes that are configured for HDFS provide NameNode and DataNode functionality. SQL Server 2019 Big Data Clusters can then leverage Isilon data either through PolyBase connectivity or HDFS Tiering providing flexible options for leveraging Isilon data in-place.

Isilon's in-place analytics approach, combined with Microsoft SQL Server, allows you to eliminate the time and resources required to replicate your Big Data into a separate infrastructure. For example, it can take over 24 hours to copy 100 TB of data over a 10 GB line. Instead, with Isilon, you can initiate data analytics projects immediately and get results in a matter of minutes.

Isilon storage solutions are designed for enterprises that want to manage data, not storage. Simple to install, manage, and scale. Nodes can be added to an Isilon cluster as needed without downtime, manual data migration, or application logic reconfiguration, thereby saving IT management resources and avoiding operational interruptions. Unlike traditional Hadoop deployments, with Isilon you also gain the flexibility to scale storage resources independently from your compute resources. And our solutions stay simple no matter how much capacity is added, how much performance is required, or how business needs change in the future.

CloudIQ

With all this conversation around analytics, lets focus now on CloudIQ, a cloud-based monitoring and storage analytics application that can be used to proactively monitors Dell EMC arrays. The value of CloudIQ centers on its ability to give users new and valuable insights into the health of the storage system.

It proactively monitors and measures overall health using intelligent, comprehensive, and predictive analytics—and that makes it easier for IT to identify storage issues quickly and accurately. These analytics (which admins can access from anywhere through a web interface or mobile app) can drive business decisions that could lower the organization's total cost of ownership associated with the array. CloudIQ delivers several key values to customers:

- Reduce Total Cost of Ownership: CloudIQ provides an easy single pane of glass from which you can monitor your systems, all from the web so you can access anytime, anywhere.

- Expedite Time to Value: Because it is deployed from the EMC Cloud, customers can simply log into their CloudIQ account and immediately access this valuable information. There is nothing to set up, no licenses, no burdens.

- Drive Business Value: CloudIQ's Proactive Health Score provides an easy way to identify and understand potential vulnerabilities in the storage environment. With these proactive and targeted guidelines, the result is a more robust and reliable storage environment, resulting in higher uptime and optimized performance and capacity.

CloudIQ is a free, cloud-based application that lets you easily monitor, analyze, and troubleshoot your storage environment from anywhere.

Summary

Storage platforms are often the key to elevating the overall performance and availability of your entire Microsoft SQL Server data landscape. Different versions of Microsoft SQL Server have will inherently have different feature sets and capabilities. Dell EMC Storage platforms will allow you to right-size your infrastructure, reduce latencies, eliminate storage silos, increase availability, and better secure your entire SQL Server investments.

Regardless of the SQL Server version, or number of versions, Dell EMC provides the hardware and software to run in the most efficient and optimized configuration. Dell EMC storage, combined with our Integrated Copy Data Management capabilities, ensures that Microsoft SQL Server databases are protected and able to be repurposed with ease. Moving forward into the world of containerized deployments and big data clusters Dell Technologies provides the most robust portfolio for ensuring success for our customers to be able to maximize performance, availability, agility and allow insights to be gathered from data in the most effective manner possible.

Resources

Find everything at: DellEMC.com/Storage-for-SQL