

Gewinnen im neuen Zeitalter des Datenmanagements

Wie Microsoft SQL Server 2019 die Art und Weise ändert, wie Sie Ihre Daten verwalten

Zusammenfassung

Im datengesteuerten Zeitalter sind Informationen genauso wertvoll wie bares Geld. Ganze Branchen werden durch neue Datenquellen, neu aufkommende Technologien und technologiebasierte Start-ups erschüttert. Die Nutzung von mehr Daten aus mehr Quellen zur schnelleren und kostengünstigeren Bereitstellung umsetzbarer Erkenntnisse ist entscheidend, um Mitbewerber überflügeln zu können.

Während die Unternehmen IT-Strategien für die Nutzung von Daten evaluieren, wird deutlich, dass die Ausführung herkömmlicher Data Warehouses und Datenbanken auf einer isolierten Infrastruktur im Zeitalter von Datenspeichern in Petabyte-Größe und Echtzeitanalysen ein unhaltbares IT-Modell darstellt.

Einen Wettbewerbsvorteil erzielen diejenigen Unternehmen, die Daten virtualisieren können, damit die Nutzer über alle Plattformen und Umgebungen hinweg schnell und einfach auf Daten zugreifen und dadurch bessere Entscheidungen in Echtzeit treffen können – ohne den Zeit- und Kostenaufwand der herkömmlichen Extraktions-, Transformations- und Lade- (ETL)-Prozesse. Zu den Anforderungen der wichtigsten Stakeholder zählen:

- Die IT muss ein neues Paradigma für das virtuelle Management der Daten vom Edge bis zum Core und vom Rechenzentrum bis zur Cloud einführen.
- Geschäfts- und IT-Verantwortliche müssen einen Weg finden, der ältere IT-Investitionen bewahrt und gleichzeitig eine Grundlage für die Zukunft schafft.

Herkömmliche relationale Datenbanken, wie z. B. Microsoft® SQL Server®, fungieren immer mehr als virtualisierter Daten-Hub, sodass Nutzer über eine einzige, einheitliche Schnittstelle auf unstrukturierte und strukturierte Daten aus dem gesamten Datenbestand zugreifen und diese verwalten können.

In diesem Whitepaper werden die Auswirkungen der Datenlast („data gravity“) und die Möglichkeiten der SQL Server 2019-Datenvirtualisierung untersucht sowie Überlegungen zum Aufbau einer agilen IT-Grundlage für das neue Zeitalter des Datenmanagements skizziert.

Oktober 2019

Inhaltsverzeichnis

Datenmanagement im datengesteuerten Zeitalter	3
Die wachsende Rolle von relationalen Datenbankmanagementplattformen	3
Daten haben Schwerkraft	4
Datenmassen erzeugen Schwerpunkte, die Anwendungen und Services anziehen	4
Datenvirtualisierung	5
Die Demokratisierung der Daten ermöglicht umfassende Analysen	5
Vereinheitlichtes Datenmanagement	6
PolyBase: Abfrage über beliebige Datentypen.....	6
Big Data-Cluster: Skalierung und Diversifizierung von Compute- und Storage-Lösungen.....	7
Neue Technologien	8
Datendienste und Infrastruktur	8
Compute-Ebene	9
Datenverarbeitung im Speicher in Echtzeit	9
Erweiterung der Verarbeitung im Speicher für größere Datenvolumen.....	9
Datenpersistenz mit nicht flüchtigem Speicher	10
Daten-Tiering und -platzierung.....	10
Storage-Ebene	10
NVMe, over Fabrics und Storage Class Memory	11
Hohe Performance für festplattenbasierte Datenvolumen.....	12
Datenplatzierung und -konsolidierung.....	12
Storage-Datendienste für Container	13
Netzwerkebene	14
Open-Networking-Technologien	14
Rechenzentrumsnetzwerk.....	15
Storage-Netzwerk	15
WAN (Wide Area Network)	15
Zusammenfassung	16
Fazit	16
Wo startet Ihr Unternehmen?	17
Dell Technologies ist bereit, Sie auf dem Weg zu unterstützen	18
Weitere Informationen	19
Weiterführende Literatur	19
Dell EMC	19
Microsoft	19
Intel	19

Datenmanagement im datengesteuerten Zeitalter

Die wachsende Rolle von relationalen Datenbankmanagementplattformen

Das datengesteuerte Zeitalter ist da. Fast jede Branche wird durch neue Datenquellen und aufkommende Technologien, die diese zu nutzen versprechen, z. B. künstliche Intelligenz (KI), maschinelles Lernen (ML) und Deep Learning (DL), erschüttert. Technologiebasierte Start-ups „stören“ auch ganze Branchen, da sie jahrzehntelange Legacy-Investitionen umgehen können und Echtzeitdatenfunktionen direkt dem Privatanwender zugänglich machen. Da Informationen so wertvoll wie bares Geld sind, sind Unternehmen im Vorteil, die mehr Daten aus mehr Quellen nutzen können, um schneller und kostengünstiger als die Mitbewerber umsetzbare Erkenntnisse bereitzustellen.

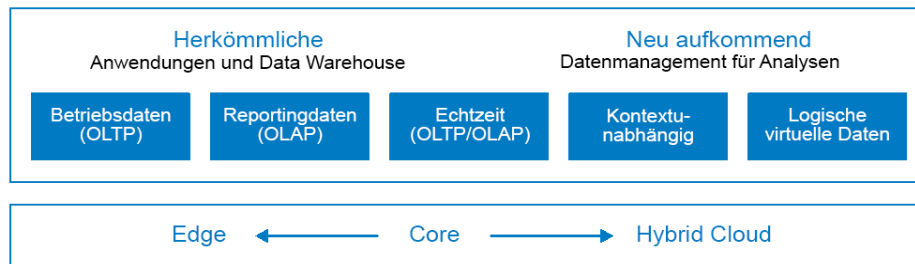
Versierte IT-Verantwortliche wissen, dass dieses neue Zeitalter einen modernen Datenmanagementansatz erfordert: einen Ansatz, der monolithische, isolierte Datenbanken durch neue Modelle für die Verwendung und gemeinsame Nutzung von Daten im gesamten Unternehmen ersetzt.

In diesem neuen Zeitalter des Datenmanagements wird SQL Server 2019 die Spielregeln ändern. Um die Auswirkungen von Microsoft SQL Server 2019 und seine Implikationen für IT-Strategien verstehen zu können, ist es jedoch wichtig, den Kontext zu verstehen, in dem sich herkömmliche relationale Datenbanken für dieses neue Zeitalter entwickelt haben.

Herkömmliche relationale Datenbanken bilden seit langem das Rückgrat für betriebliche Daten (Online-Transaktionsverarbeitung [OLTP]) und herkömmliche Data Warehouses (Online-Analyseverarbeitung [OLAP]) und unterstützen Reporting und einige Analysefunktionen. Da die Unternehmen jedoch nach Möglichkeiten suchen, die in ihren Daten verborgenen Erkenntnisse – über verschiedene Plattformen und Umgebungen hinweg – zu erschließen, wurden die Datenbankmanagementlösungen weiterentwickelt, um zusätzliche Analysen zu ermöglichen.

Relationale Datenbankmanagementsysteme (RDMS)

Herkömmliche und neue Anwendungsfälle



Diese wichtige Änderung der Datenbankmanagementstrategie durch große Anbieter hat Gartner dazu veranlasst, eine neue Kategorie mit dem Namen „Data Management Solutions for Analytics“ (DMSA, Datenbankmanagementlösungen für Analysen) zu definieren, in der Microsoft als führend anerkannt wurde.¹ Nachfolgend sind von Gartner definierte Anwendungsfälle für die Arbeit mit verteilten Daten aufgeführt.²

- **Herkömmliches Data Warehouse.** Dieser Anwendungsfall umfasst das Management strukturierter historischer Daten aus mehreren Quellen. Die Daten werden hauptsächlich mittels Bulk- und Batch-Loading geladen. Das herkömmliche Data Warehouse kann große Datenmengen verwalten und wird hauptsächlich für standardmäßiges Reporting und für Dashboards verwendet. In geringerem Umfang wird es auch für Freiform-Ad-hoc-Abfragen und Mining oder betriebliche Abfragen genutzt.
- **Echtzeit-Data-Warehouse.** Bei diesem Anwendungsfall werden Analyseanwendungen um eine Echtzeitkomponente ergänzt, damit sich die Latenzzeit zwischen dem Zeitpunkt der Datengenerierung und der Analyse verkürzt. Bei diesem Anwendungsfall werden in erster Linie strukturierte Daten verwaltet, die kontinuierlich über Mikro-Batching und/oder Streaming-Aufnahme zur Entscheidungsunterstützung in Echtzeit sowie zur Unterstützung von in Anwendungen eingebetteten Analysen, von Echtzeit-Data-Warehousing und von Betriebsdatenspeichern geladen werden.
- **Logisches Data Warehouse.** Bei diesem Anwendungsfall werden die Datenvielfalt und die Datenmengen sowohl für strukturierte Daten als auch für Datentypen anderer Inhalte verwaltet, wobei DMSA als logische Schicht für eine Vielzahl von Datenquellen dient. Neben strukturierten Daten aus transaktionalen Anwendungen umfasst dieser Anwendungsfall auch Datentypen anderer Inhalte, z. B. Maschinendaten, Textdokumente, Bilder und Videos.

¹ Microsoft, „Gartner nennt Microsoft einen Marktführer im Magic Quadrant for Data Management Solutions for Analytics (DMSA)“, Januar 2019.

² Gartner, „Magic Quadrant for Data Management Solutions for Analytics“, Januar 2019.

- **Kontextunabhängig.** Dieser Anwendungsfall ermöglicht die Erforschung neuer Datenwerte, Datenformatvarianten und Beziehungen. Er unterstützt Such-, Diagramm- und andere Funktionen zum Aufdecken neuer Informationsmodelle. Dieser Anwendungsfall wird hauptsächlich für Freiformabfragen zur Unterstützung von Prognosen, vorausschauender Modellierung oder anderen Mining-Formaten sowie für Abfragen verwendet, die mehrere Datentypen und -quellen unterstützen.

Der Entwurf und die Entwicklung der besten Lösung für Ihren Datenbestand – vom Edge bis zum Core und von On-Premise bis zur Hybrid Cloud – erfordert eine Kombination aus dem richtigen Fachwissen und der richtigen Infrastruktur. Es handelt sich nicht wie früher um eine Einheitslösung, sondern um einen verbrauchsorientierten, flexiblen und skalierbaren Ansatz, der auf aktuellen Investitionen aufbaut und Sie gleichzeitig auf eine erfolgreiche digitale Zukunft vorbereitet.

In diesem Whitepaper werden wir Ihre SQL Server-Reise in Verbindung mit der Vision von Dell EMC für das Management Ihres Datenkapitals als Teil Ihres gesamten Datenbestands untersuchen. Dell Technologies ist einzigartig positioniert, um Lösungen vom Edge über den Core bis zur Hybrid Cloud bereitzustellen, um die in Microsoft SQL 2019 verfügbaren Tools zu unterstützen, zu schützen und wiederzuverwenden.

Daten haben Schwerkraft

Datenmassen erzeugen Schwerpunkte, die Anwendungen und Services anziehen

Es ist kein Geheimnis, dass das Datenvolumen groß ist – und immer größer wird. Einigen Schätzungen zufolge erstellen wir jeden Tag 2,5 Trillionen Byte Daten.³ Im Zuge der steigenden Cloud-Akzeptanz, der wachsenden Zahl von Mobilgeräten und der Ausbreitung des IoT (Internet der Dinge) kommt es zu einem gewaltigen Wachstum der Datenmengen außerhalb der Grenzen des herkömmlichen Rechenzentrums.

Die Datenmengen sind mittlerweile so groß und so stark verteilt, dass der Ort, an dem die Daten erzeugt und genutzt werden, zu einem entscheidenden Faktor für die Beschleunigung der Workloads und der Entscheidungsfindung geworden ist. Je größer die Datenmenge, desto schwieriger ist es, sie zu verschieben. Daher werden Anwendungen, Services und sogar andere Daten von diesen Daten angezogen und umkreisen sie wie einen Schwerpunkt: Genauso wie Sternstaubwolken, die sich einst zu Sternen und Planeten verdichteten und nun eine Gravitationskraft („gravity“) auf die Objekte in ihrer Umgebung ausüben, gewinnen Daten an Masse und ziehen Anwendungen und Services an.

Durch die Platzierung von Anwendungen und Services in der Nähe der Daten können Sie die Latenz reduzieren und die Leistung verbessern. Analysen werden schneller durchgeführt, damit Sie schneller bessere Entscheidungen treffen können. Die Antwortzeiten sind kürzer, wodurch sich die Kundenzufriedenheit erhöht. Kurz gesagt: Die Daten müssen aufgenommen, aggregiert, analysiert und dort eingesetzt werden, wo sie sich befinden.

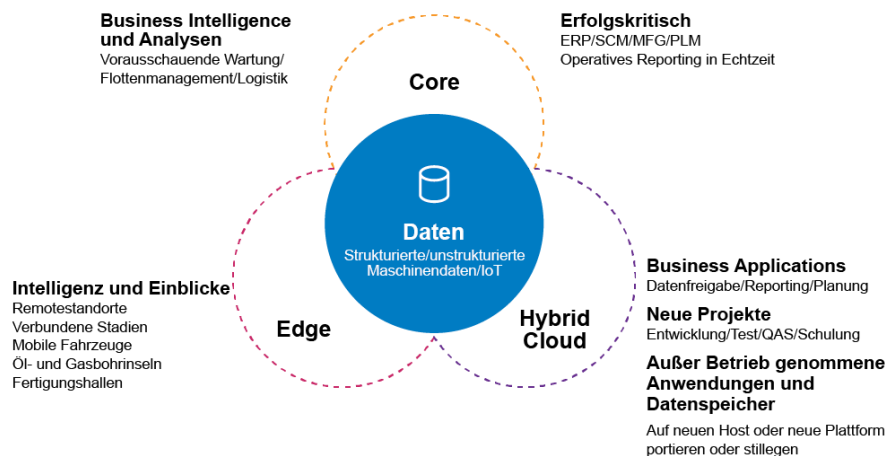
Beispielsweise muss ein autonomes Fahrzeug mehrere Datenpunkte (von Kameras, Radar und Sonar) aufnehmen und in Echtzeit verarbeiten. Wenn sich ein Objekt in den Fahrweg des Wagens bewegt, bleibt keine Zeit, die Daten in die Cloud oder ins Rechenzentrum zu senden und auf eine Antwort zu warten: Die Analysen müssen zu den Daten am Edge kommen. Wenn die Daten in die und aus der Cloud gestreamt werden (beispielsweise für Mobilfunkdienste), müssen sich die Anwendungen und Analysen in der Cloud befinden, damit die Daten in Echtzeit verarbeitet werden können. Ebenso können die Daten, die Anwendungen für mobile und Remotenutzer versorgen, sicher im Rechenzentrum bleiben, wodurch der Schutz für vertrauliche personenbezogene Daten und Finanzdaten verbessert wird.

Während sich die modernen Unternehmen an die digitale Wirtschaft anpassen, muss die IT mit einer Datenmanagementstrategie darauf vorbereitet sein, die es dem Unternehmen ermöglicht, Daten zu nutzen, die sich in verteilten hybriden Umgebungen vom Edge über den Core bis zur Cloud befinden.

³ Forbes, „[How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read](#)“, Mai 2018.

Auswirkungen der Datenlast

Anwendungen | Datenplatzierung | Datenmanagement



Durch die Übernahme einer IT-Strategie, die den Prinzipien der Datenlast („data gravity“) Rechnung trägt, können Sie Ihre Daten unabhängig davon nutzen, wo sie sich befinden. An dieser Stelle kommt die Datenvirtualisierung ins Spiel. Sie wird die Welt, wie wir sie kennen, verändern.

Datenvirtualisierung

Die Demokratisierung der Daten ermöglicht umfassende Analysen

Mit „Datenvirtualisierung“ ist die Abstraktion von Daten aus unterschiedlichen Quellen, Standorten und Formaten – ohne diese Daten zu kopieren oder zu verschieben – in einer einzigen Ebene gemeint, die es Nutzern erlaubt, sie in Echtzeit über eine einzige, einheitliche Schnittstelle abzufragen. Dadurch können auch sensible Daten von einem einzigen Standort aus kontrolliert werden. Da es nicht mehr notwendig ist, mehrere Kopien der Daten zu erstellen, werden auch die Kosten für Speicherung und Management minimiert.

Einige der wichtigsten Triebkräfte der Datenvirtualisierung sind die Herausforderungen, die herkömmliche Data Warehouses für Business Intelligence (BI) mit sich bringen. Im Kern handelt es sich um die Anforderung, Daten zu kopieren und in die Reporting-Plattform des Data Warehouse zu laden. In der Vergangenheit wurde die ETL-Verarbeitung – oder „Daten-Pipelining“ – für die Datentransformation eingesetzt. Dieser Prozess hat Herausforderungen mit sich gebracht, die nicht ignoriert werden können, während sich die Unternehmen bemühen, besser und in Echtzeit auf ihre Kunden und Partner und auf die Marktdynamik zu reagieren. Zu den größeren Herausforderungen zählen:

- **Datenlatenz.** ETL-Pipelines führen zu einer inhärenten Verzögerung. Eine aktuelle Studie hat gezeigt, dass mehr als 80 % der Datenvolumina, die durch ETL-Pipelines bereitgestellt werden, zwischen zwei und sieben Tage alt sind, wenn sie ein Analysesystem erreichen. Ganze 75 % der Unternehmen gaben an, dass Verzögerungen bei der Datenverarbeitung Geschäftschancen gehemmt haben.³
- **Big Data und Analysen.** Durch die Ausdehnung des IoT generieren vernetzte Geräte und Mitarbeiter Datenmengen, die die Speicherkapazität herkömmlicher Datenbanksysteme überschreiten. Schätzungen zufolge wurden 90 % des gesamten Datenbestands der Welt in den letzten zwei Jahren mit einer Geschwindigkeit von $2,5 \times 10^{18}$ Byte Datenmenge pro Tag generiert.⁴ Diese neue Art von Daten wird häufig in Formaten verwendet, die nicht für die Speicherung in relationalen Datenbanktabellen oder für Abfragen mit relationaler Abfragesemantik geeignet sind.
- **Big Data braucht „Kontext“.** Viele IoT-Datenvolumina sind unstrukturiert und überwältigend in Volumen, Geschwindigkeit und Vielfalt. Zudem sind sie ohne Unternehmenskontext wertlos. Die IT muss die Brücke bauen, damit geschäftliche/betriebliche Daten, die in relationalen Datenbanken gespeichert sind, mit IoT-Daten zusammenkommen, um Echtzeit-BI und Antwortzeiten zu ermöglichen, die zu selbst lernenden intelligenten Anwendungen und Geschäftsprozessen führen, die KI und ML nutzen.

⁴ Technisches Whitepaper von Microsoft, [Microsoft SQL Server 2019 Big Data-Cluster](#), September 2018.

Vereinheitlichtes Datenmanagement

Echtzeitdatenzugriff und Transparenz sind für Unternehmen immer wichtiger, um einen Wettbewerbsvorteil zu erhalten. Gleichzeitig überschreiten neue Anwendungen und neue Anwendungsfälle, die z. B. durch KI und IoT ermöglicht werden, die Grenzen herkömmlicher isolierter Rechenzentren und Prozesse wie ETL. Tatsächlich vertauscht die IT das „T“ und das „L“, um die Datenverschiebung zu beschleunigen, wobei die Übertragung der Daten in das Ziel optimiert und dem Zielsystem die Übersetzung überlassen wird. Transformation wird jetzt von der Datenplattform gesteuert, auf der sich die Daten befinden.

Angesichts der zunehmenden Akzeptanz der Datenvirtualisierung ist der Bedarf an Datenmanagementplattformen und -infrastrukturen gewachsen, die diese virtuelle Datenschicht mit einer einheitlichen Sicht verwalten können, die mehrere Anwendungen, Nutzer, Datenspeicher und Standorte überbrückt. SQL Server 2019 ist ein Datenbankmanagementsystem, dessen Funktionen als Plattform für die Vereinheitlichung von Daten weiter ausgebaut werden.

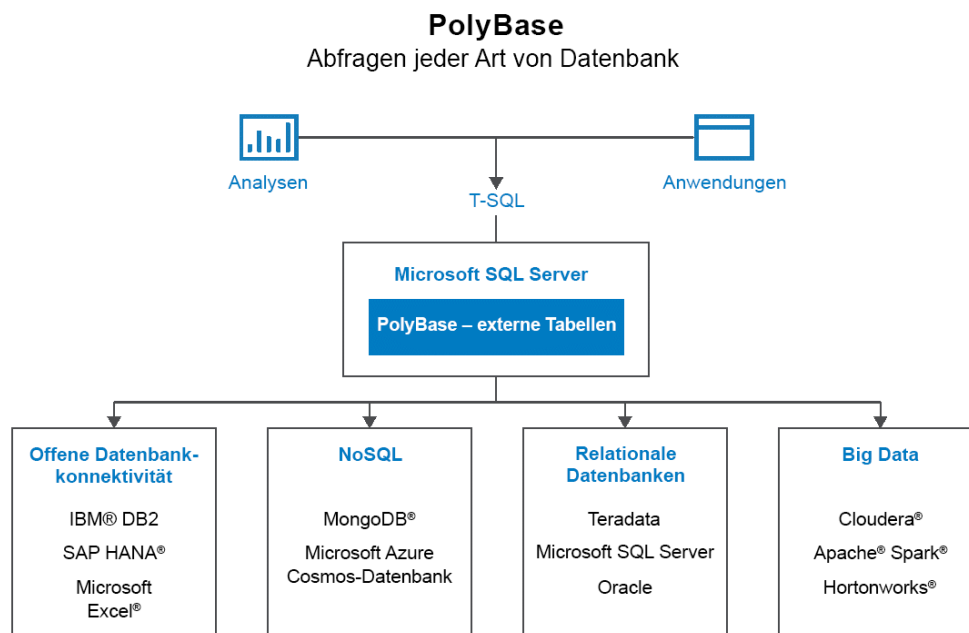
SQL Server 2019 ist ein Hub für Daten und kann Daten aus Oracle®, Teradata und Hadoop® auf eine sichere und sehr leistungsstarke Weise abfragen – ohne sie zu verschieben oder zu kopieren. Zu den wichtigsten Funktionen von SQL Server 2019 zur Vereinheitlichung von virtuellen Geschäftsdaten und Betriebsdaten, die in relationalen Datenbanken gespeichert werden, mit IoT-Daten für BI in Echtzeit sowie eingebettete KI und ML gehören:

PolyBase: Abfrage über beliebige Datentypen

PolyBase ist eine Technologie, die sowohl auf nicht relationale als auch auf relationale Daten, die sich an unterschiedlichen Standorten befinden, zugreift und diese kombiniert – und das alles innerhalb von SQL Server. Sie ermöglicht es Anwendungen und Nutzern, eine Vielzahl von Datenspeichern abzufragen, einschließlich solcher, die offene Datenbankkonnektivität, NoSQL, relationale Datenbanken und Big Data-Speicher in mit Hadoop Distributed File System (HDFS) kompatiblen Distributionen und Dateisystemen unterstützen.

PolyBase verschiebt einige Berechnungen nach unten, um die Leistung zu optimieren. Operationen wie Projektionen, Prädikate, Aggregate, Grenzwerte und homogene Vereinigungen werden an das Quellsystem übergeben, damit der Abfrageoptimierer in jedem der Quellsysteme genutzt werden kann. Es werden nur gefilterte Ergebnisse zurückgegeben, wodurch die Menge der zu übertragenden Daten reduziert und somit die Leistung gesteigert wird.

SQL Server 2019 erweitert die Funktionen von PolyBase durch neue Connectors, um externe Tabellen zu erstellen, die mit einer Vielzahl von Datenspeichern verknüpft werden können, z. B. SQL Server, Azure® SQL-Datenbank, Azure SQL Data Warehouse, Oracle, Teradata, MongoDB, Azure Cosmos DB oder über einen generischen ODBC-Treiber mit jeder ODBC-kompatiblen Datenquelle (Open Database Connectivity).



Quelle: Microsoft SQL Server 2019 Big Data-Cluster – Technisches Whitepaper

Wenn Teams problemlos Daten aus verschiedenen Quellen unabhängig von Standort und Typ nutzen können, haben sie die Grenzen, die eine umfassende Datenanalyse und eine beschleunigte Entscheidungsfindung und Wertschöpfung in ihrem Unternehmen verhindern, effektiv beseitigt.

Darüber hinaus können Entwickler in vielen Fällen die Attribute der T-SQL-Prozedur-Ergebnismenge, die an die Anwendungen und Analysen zurückgegeben werden, konsistent halten, selbst wenn die Daten von einer neuen Quelle über externe PolyBase-Tabellen zurückgegeben werden.

Big Data-Cluster: Skalierung und Diversifizierung von Compute- und Storage-Lösungen

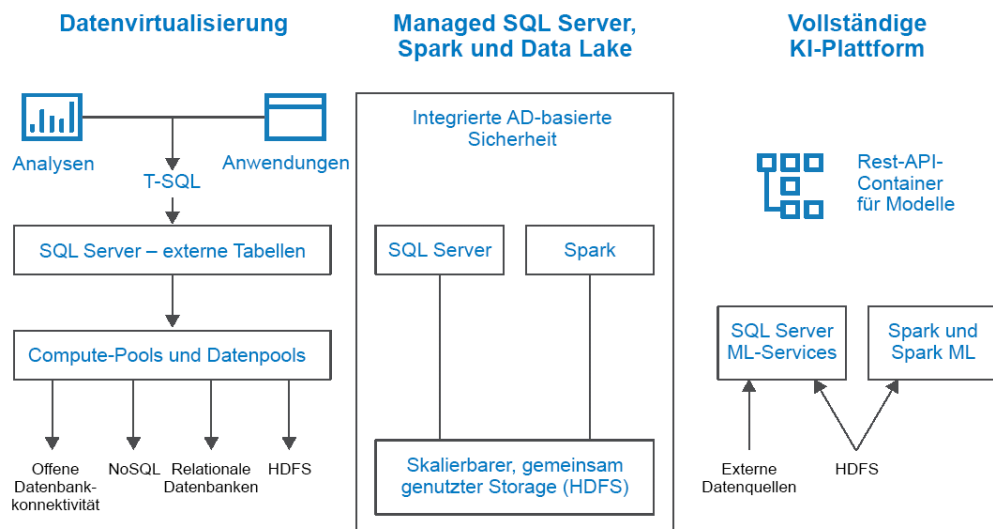
SQL Server 2019 Big-Data-Cluster nutzen PolyBase zur Verbesserung der Datenvirtualisierungserfahrung durch eine schnellere und sicherere Integration von SQL Server in Apache HDFS- und Spark-Systeme.

SQL Server Big-Data-Cluster integrieren SQL Server und branchenübliche Big-Data-Tools in einem einzigen Paket, das einfach bereitzustellen und verwalten ist. Sie bieten drei Hauptfunktionen:

- **Datenvirtualisierung** ermöglicht die Kombination von Daten aus mehreren Quellen, ohne diese zu verschieben oder zu replizieren. Die Leistung kann durch die horizontale Skalierung von Compute und Caching gesteigert werden.
- Mit **Managed SQL Server, Spark und Data Lake** kann man große Datenmengen in einem Data Lake speichern und über SQL oder Spark darauf zugreifen. Darin enthalten sind Managementservices und integrierte Sicherheit, um das Management zu vereinfachen.
- **Vollständige KI-Plattform.** Integrierte Daten aus vielen Quellen können problemlos in Trainingsmodelle eingespeist werden. Sie können Daten einlesen, aufbereiten, dann Modelle in einem einzigen System trainieren, speichern und operationalisieren.

Microsoft SQL Server 2019

Big Data-Cluster mit KI zur Bereitstellung intelligenter Anwendungen

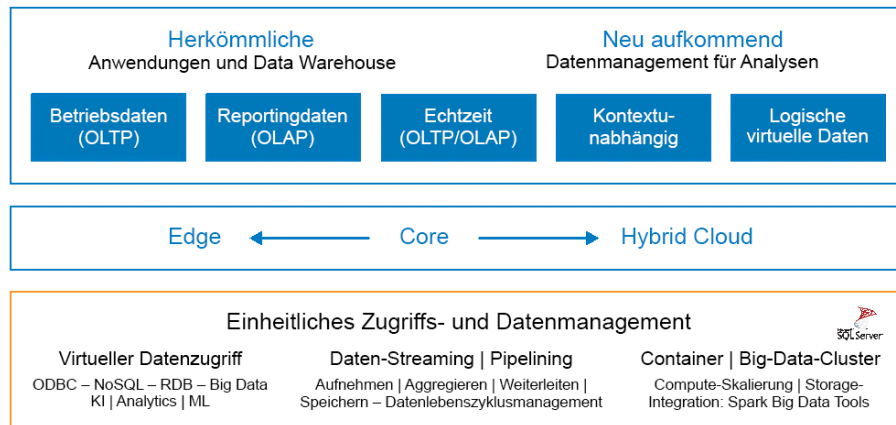


Quelle: Microsoft SQL Server 2019 Big Data-Cluster – Technisches Whitepaper

Die Unterstützung von Docker-Containern, Linux® und PolyBase in SQL Server stellt die Voraussetzungen für SQL Server 2019 Big-Data-Cluster dar. Big-Data-Cluster können Container in Kubernetes vollständig für die Bereitstellung von Anwendungen nutzen und stellen damit eine vorhersehbare, schnelle und flexibel skalierbare Bereitstellung sicher, unabhängig davon, wo der Kubernetes-Cluster bereitgestellt wird. Da Unternehmen Container für Anwendungsfälle mit statusabhängigen Anwendungen nutzen möchten, wird der Bedarf an Datenpersistenz, Schutz und Management eine zentralere Rolle spielen. Sie müssen Storage-Lösungen in Betracht ziehen, die die für die Unterstützung dieser Plattformen erforderliche Datenpersistenz und die erforderlichen Services bieten.

Relationale Datenbankmanagementsysteme (RDMS)

Herkömmliche und neue Anwendungsfälle



Die Datenvirtualisierung hat starke Auswirkungen. Da die Daten nicht für die am besten ausgebildeten Data Scientists zugänglich sind, können sie nun in mehrere Berichte, Dashboards und Anwendungen aufgenommen werden, wodurch sich ihr Wert multipliziert.

Es ist klar, dass die Zunahme der Datenvirtualisierung und die Einführung neuer Funktionen für eine einheitliche Datenmanagementplattform eine neue Strategie für Datendienste und Infrastrukturen erfordern, die SQL Server 2019 ergänzen. Im nächsten Abschnitt wird die Dell EMC Strategie für Datendienste und Infrastrukturen, die SQL Server 2019 ergänzen, erläutert.

Neue Technologien

Datendienste und Infrastruktur

Es fühlt sich vielleicht so an, als wären Sie schon einmal hier gewesen. Die IT-Landschaft hat sich deutlich verändert und jetzt müssen Sie sich wieder neu ausrichten, um wettbewerbsfähig zu bleiben. In der Vergangenheit bedeutete dies, dass relativ neue Investitionen ausgetauscht und ersetzt werden mussten. Das ist, gelinde gesagt, kein ideales Szenario.

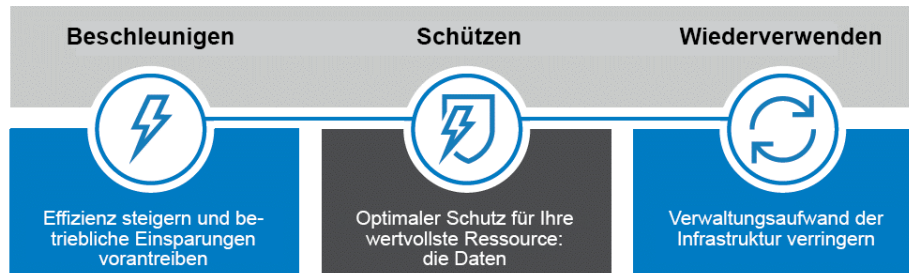
Diesmal ist es anders. SQL Server 2019 wurde für die Integration des Datenbestands entwickelt, sodass man sich von einem Entweder-oder-Denken hin zu einem strategischeren verbindenden Ansatz entwickeln kann. Sie können diese einheitliche Ebene für virtualisierte Daten On-Premises und in der Cloud auf herkömmlichen dreischichtigen Architekturen mit All-Flash-Arrays und Software Defined Storage sowie konvergenten Infrastrukturen und hyperkonvergenten Infrastrukturen ... und so weiter nutzen.

Natürlich sollten Sie in Betracht ziehen, neueste Infrastrukturtechnologien bereitzustellen, um alle Leistungs- und Agilitätsfunktionen von SQL Server 2019 nutzen und von der Arbeit mit virtuellen Daten profitieren zu können. Die neuen Infrastrukturtechnologien unterstützen Sie bei folgendem:

- **Beschleunigen** und Gewährleisten der Zukunftssicherheit Ihrer Umgebung, während Sie Ihre SQL-Infrastruktur vollständig modernisieren. Eine moderne Sicht auf Storage nutzt Storage Class Memory, persistenten Speicher wie NVDIMM und andere Arbeitsspeichertechnologien, um maximale Wirkung zu erzielen.
- **Schützen** Ihrer Datenbanken durch branchenführende Backups, Replikationen, Snapshots, Ausfallsicherheit und selbst erstellten deduplizierten Kopien.
- **Wiederverwenden** von Snapshots für die betriebliche Recovery, für die Neuzuweisung von Entwicklungs- und Testzwecke sowie für Pipelines für die fortlaufende Integration/kontinuierliche Bereitstellung (CI/CD). Sie könnten z. B. eine Datenbank aus der Produktion auf Storage-Ebene kopieren und die Daten bearbeiten, damit sie zu einer Quelle für externe PolyBase-Tabellen werden können, ohne dass die Produktdatenbank beeinträchtigt wird.

Durch die Ausrichtung dieser Säulen wird die Effizienz und Konsistenz eines einheitlichen Ansatzes für das Management des Datenbestands gesteigert. Durch die Kombination mit einer starken, konsistenten, hochleistungsfähigen Architektur, die die Datenbankplattform unterstützt, kann Ihr IT-Team die Modernisierung meistern.

Einrichten einer dynamischen IT-Grundlage



Nehmen wir uns einen Moment Zeit, um neue Infrastrukturtechnologien (Compute, Storage, Netzwerke) und Anwendungsfälle für das Datenmanagement mit SQL Server 2019 zu erforschen.

Compute-Ebene

Ab der Compute-Ebene ist die Datenverarbeitung im Speicher in Echtzeit, die DRAM nutzt, zentral für die Beschleunigung leistungsabhängiger Anwendungsfälle im Zusammenhang mit Microsoft SQL. Beispielsweise nutzt SQL Server die Datenverarbeitung im Speicher bei Data Analytics in Echtzeit und kontextunabhängigen Analysen zu folgenden Zwecken:

- Beschleunigung von Echtzeitanalysen und Reporting durch Optimierung der Leistung von OLTP und Aggregation von OLAP-Daten für BI, einschließlich Datenaufnahme, Datenlast und Szenarien mit kurzlebigen Daten;
- Unterstützung bei der Erforschung neuer Datenwerte, Datenformvarianten und Beziehungen zur Durchführung von Freiformabfragen zur Unterstützung von Prognosen, vorausschauender Modellierung und anderen Mining-Stilen sowie für Abfragen mehrerer Datentypen und -quellen; sowie
- Verarbeitung größerer Datensätze für Strukturen und andere Inhaltstypen, z. B. Maschinendaten, Textdokumente, Bilder und Videos.

Datenverarbeitung im Speicher in Echtzeit

Relationale Datenbanken, wie z. B. SQL Server, haben die Datenverarbeitung im Speicher für OLTP und OLAP übernommen, um die Leistung von Transaktionsverarbeitung, Datenaufnahme, Datenlast und Szenarien mit kurzlebigen Daten zu optimieren.

Die Verarbeitung im Speicher ist zwar in SQL Server nicht neu, jedoch werden die Datenbanken immer größer und umfassen immer häufiger mehrere Terabyte. Die gute Nachricht ist, dass Server, z. B. Dell EMC Server mit zwei und vier Sockeln und Intel® Prozessoren ohne Unterbrechungen auf bis zu 6 TB Arbeitsspeicher erweitert werden können und damit Skalierbarkeit bieten, wenn SQL Server-Bereitstellungen wachsen.

Die Server unterstützen zwar größere speicheroptimierte Datenbanken für vertikal skalierbare und horizontal skalierbare Systemlandschaften, doch die Datenpersistenz wird traditionell in der Storage-Schicht sichergestellt, um Datenverluste bei geplanten und ungeplanten Systemausfällen zu vermeiden.

Im Jahr 2019 hat Dell EMC PowerEdge-Server mit persistentem Intel Optane™ DC-Speicher auf den Markt gebracht. Intel Optane DC ist eine innovative Arbeitsspeichertechnologie, die eine einzigartige Kombination von erschwinglicher, großer Kapazität und Unterstützung für Datenpersistenz bietet. Dies ist eine neue Klasse von Arbeitsspeicher und Storage, die die Gewinnung von Erkenntnissen aus den enormen Datenmengen, die Unternehmen heute verwalten, beschleunigen kann.

Erweiterung der Verarbeitung im Speicher für größere Datenvolumen

Persistenter Intel Optane DC-Speicher kann mit herkömmlichem DRAM als Cache konfiguriert werden, sodass er transparent in das Speichersubsystem integriert wird und wie DRAM erscheint, ohne dass Änderungen am Betriebssystem (BS) oder an Anwendungen erforderlich sind.

Dies bietet einige Vorteile. Erstens können Sie den Arbeitsspeicher der heutigen Dell EMC PowerEdge-Server mit vier Sockeln von 6 TB DRAM auf 12 TB nutzbaren System Speicher (3 TB pro Sockel) verdoppeln.⁵ Dies ermöglicht eine Konsolidierung durch die Reduzierung der Anzahl von Systemknoten. Zweitens kann teurer DRAM-Arbeitsspeicher durch kostengünstigere persistente Intel Optane DC-Speichermodule ersetzt werden.

⁵ Intel, „[The Challenge of Keeping Up With Data](#)“, April 2019.

Datenpersistenz mit nicht flüchtigem Speicher

Den Anwendungen, die für die Ausführung im Speicher (DRAM) entwickelt wurden, müssen die Datenbanktabellen und Daten zur Ausführung der OLTP- und OLAP-Prozesse in der Regel im Arbeitsspeicher zur Verfügung stehen. Bei einer geplanten oder ungeplanten Systemausfallzeit gehen alle Daten, die im Arbeitsspeicher verarbeitet werden, verloren. Nach dem Neustart müssen die Datenbanktabellen und Daten in den Arbeitsspeicher geladen werden, bevor die Anwendungen neu gestartet werden kann.

Was die Storage-Leistung betrifft, ist der Arbeitsspeicher am schnellsten, da er sich näher am Prozessor am Speicherbus befindet. Größere Datenbanktabellen, die in persistentem Optane DC-Speicher gespeichert werden, stehen für das schnellere Laden von Daten nach einem Ausfall zur Verfügung.

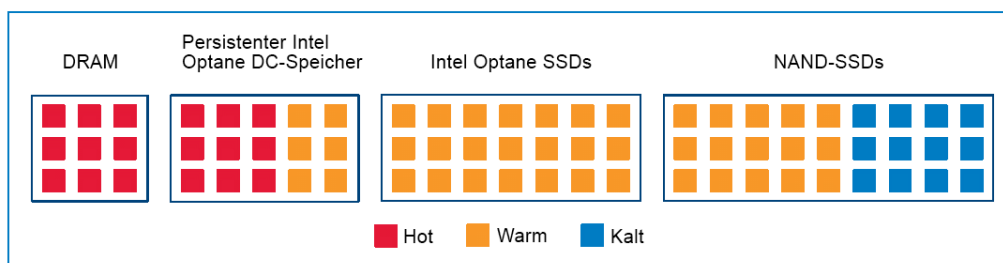
Der Vorteil der persistenten Speicherung von größeren Tabellen und Daten im persistenten Intel Optane DC-Speicher besteht im schnelleren Laden von Daten und besseren Recovery Time Objectives (RTO). Beispiel: Tests von Dell EMC, in denen Microsoft SQL 2019 in einer Vorschauumgebung ausgeführt und persistenter Optane DC-Speicher genutzt wurde, haben gezeigt, dass Nutzer eine bis zu 2,7-fache Verbesserung der SQL-Datenbankleistung feststellen können.⁶

Daten-Tiering und -platzierung

Die Datennutzung ändert sich im Laufe der Zeit. Es kann vorkommen, dass eine Gruppe von Daten anfänglich häufig verwendet wird, die Häufigkeit jedoch im Laufe der Zeit abnimmt, wenn die Daten älter werden. Dies ist keine neue Offenbarung. Jedoch sollten Unternehmen in Anbetracht dessen, an wie vielen Orten ihre Daten gespeichert werden – vom Edge über den Core bis zur Hybrid Cloud – ihre Strategien zum Datenlebenszyklusmanagement neu bewerten. Dazu gehören Storage-Strategien für das Tiering in „hot“, „warm“ und „cold“ Daten und die Verlagerung von eingefrorenen Daten in die Cloud wie Microsoft Azure.

Daten-Tiering umfasst das Arbeiten über verschiedene Arbeitsspeicher- und Storage-Klassen. Wir werden diesen Anwendungsfall im nächsten Abschnitt näher untersuchen.

Speicherpools und Daten-Tiering



Diese Abbildung veranschaulicht Pooling und Daten-Tiering mit DRAM, persistentem Intel Optane DC-Speicher und Intel Optane DC-SSDs mit „hot“ Daten in Rot, „warm“ Daten in Orange und „cold“ Daten in Blau.

Storage-Ebene

Die Datenvirtualisierung versetzt Unternehmen in die Lage, überall auf Daten zuzugreifen, gleichgültig, wo sie sich befinden. Storage spielt eine wichtige Rolle bei der Unterstützung von Unternehmen, den Wert ihrer Daten zu erschließen und ihr Datenkapital zu maximieren. Es muss darauf geachtet werden, wo sich die Daten befinden, wie sie verwaltet und wie sie geschützt werden. Storage war – wie bereits besprochen – schon immer eine treibende Kraft für die Konzepte von Beschleunigung, Schutz und Wiederverwendung. Neuentwicklungen der Storage-Technologie unterstützen Unternehmen weiter dabei, von der Arbeit mit virtuellen Daten zu profitieren.

Um den Wert von Geschäftsdaten erschließen zu können, muss darauf geachtet werden, wo sich die Daten befinden, wie sie verwaltet und wie sie geschützt werden.

Wir stellen in Unternehmen einen Wandel von einem Cloud-zentrierten Ansatz zu einer datenzentrierten Denkweise fest. Auf diese Weise lässt sich sicherstellen, dass sich Ihr wertvollstes Gut zum richtigen Zeitpunkt am richtigen Ort befindet und über das richtige Service Level Agreement (SLA) verfügt. Organisatorische Anforderungen und Geschäftsmodelle sollten für die Entscheidung maßgeblich sein, wo sich die Daten befinden, wobei gleichzeitig die Konsolidierungsstrategien zu berücksichtigen sind, die zu einer Reduzierung der Administrationsaufgaben und Leistungssteigerung beitragen können.

⁶ StorageReview, „[Dell EMC PowerEdge Servers Refreshed with New Intel CPUs and Optane DC Persistent Memory](#)“, April 2019.

SQL Server unterstützt beispielsweise die Speicherung von Tabellen und Daten sowohl in In-Memory-Datenbank als auch auf Festplatten. Allerdings basiert die Wahl des Speicherorts der Daten auf Faktoren wie dem Anwendungsprofil, den Datentypen und den zur Unterstützung der Nutzer erforderlichen IT-Services.

Microsoft bietet zwar einige Hinweise für die Platzierung von Tabellen und Daten, um die IT-Planung zu vereinfachen, jedoch erfordert die Datenplatzierung in der Regel einiges Nachdenken, wobei die SLA-Anforderungen hinsichtlich Leistung und die Verfügbarkeit zu berücksichtigen sind, einschließlich des Datenlebenszyklusmanagements im Vergleich zu den IT-Kosten für die Bereitstellung und das Management von Service und Daten.

NVMe, over Fabrics und Storage Class Memory

Es ist wichtig, sich neue Storage-Technologien anzuschauen, die bei der Planung der IT-Grundlage für Anwendungsfälle und Services für das Datenmanagement mit SQL Server 2019 zur Verfügung stehen. All-Flash-Storage ermöglicht beispielsweise hohe Performance, Durchsatz und Datenreplikation für Anwendungsfälle, wie z. B. Systemkopie/-Aktualisierung für Entwicklungs-/Testzwecke und betriebliche Recovery mit neuen Verfahren, wie fortlaufende Integration und Bereitstellung.

Im Hinblick auf die Leistung hat Flash-Storage bereits durch die Bereitstellung einer hohen Performance mit einer Latenz von einer Millisekunde oder weniger die Verarbeitung von OLTP- und OLAP-Workloads grundlegend verändert. Für viele Kunden bringt die Modernisierung auf ein All-Flash-Array bedeutend mehr IO und Bandbreite mit sich. Dies führt zu einem weniger aufwendigen Systemmanagement, weniger Anrufe bei Support-Desks und massiven Verbesserungen beim Zeitaufwand für die Erstellung von Berichten.

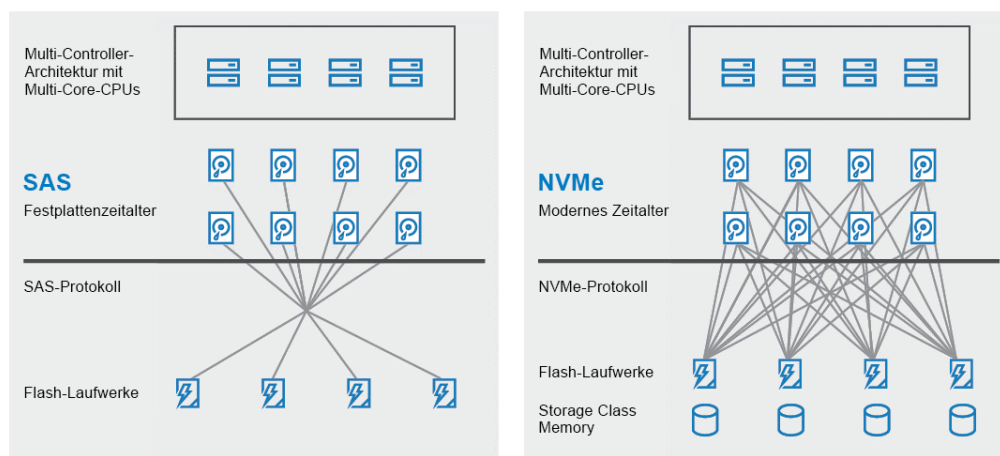
Bei den neuen BI-, Echtzeitanalysen und Big Data-Anwendungsfällen mit SQL Server 2019 ist Flash erst der Anfang. Neue Technologien wie Non-volatile Memory Express (NVMe), NVMe over Fabrics (NVMe-oF) und Storage Class Memory (SCM) steigern den Wert von Flash-SSDs, die zu grundlegenden Storage-Plattformen zur Unterstützung dieser Anwendungsfälle werden.

Nehmen wir uns einen Moment Zeit, um zu betrachten, wie sich diese Technologien entwickelt haben und wie sie sich auf die Leistung auswirken.

NVMe gibt es seit 2012. NVMe wurde entwickelt, um die Vorteile der Parallelität moderner CPUs und SSDs zu nutzen. Es ermöglicht eine schnellere Datenübertragung zum Host-Controller-System, z. B. 32 GB/s für NVMe im Vergleich zu 12GB/s für SAS 3.0 und ist damit fast dreimal so schnell.

Maximierung der Leistung einer Multi-Controller-Architektur

Reduzierung der Antwortzeit von unter Millisekunden auf Mikrosekunden



NVMe durchbricht die Engpässe, die auftreten, wenn der schnelle Flash-basierte SSD-Speicher mit Legacy-Datentransporttechnologien kollidiert. Vor allem öffnet diese Technologie die Tür zur nächsten Medienrevolution mit SCM.

Beispielsweise bietet Dell EMC Storage mit einer Multi-Controller-Architektur und durchgängigem NVMe, um die Leistung mit modernen Medien, einschließlich Flash-SSDs und SCM, zu maximieren. Sie können heute in All-Flash-Storage von Dell EMC investieren und morgen unterbrechungsfrei NVMe- und SCM-Laufwerke hinzufügen.⁷

NVMe-OF, das im Jahr 2016 vorgestellt wurde, ist eine Weiterentwicklung mit einem Transportprotokoll, das einen Hostcomputer, z. B. einen Dell EMC Server, auf dem Microsoft SQL Server ausgeführt wird, und ein Ziel-Solid-State-Speichergerät oder -System ermöglicht, z. B. Dell EMC Storage über Fibre Channel (FC).

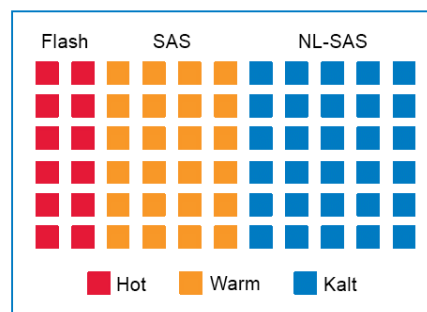
SCM, auch als persistenter Speicher bezeichnet, ist möglicherweise die revolutionärste neue Storage-Technologie des Jahrzehnts. SCM ist zwar etwas langsamer als DRAM, hat jedoch den Vorteil der Persistenz. Das heißt, dass der Speicherinhalt wie bei herkömmlichem Storage nicht verloren geht, wenn das Gerät ausgeschaltet wird.

Der entscheidende Punkt ist, dass die Ausführung von Datenbanken im Speicher (DRAM) schneller ist als im festplattenbasierten Storage. Storage-Plattformen haben neue Technologien eingeführt, um die Leistungsanforderungen von Analysen und Big-Data-Szenarien wie Transaktionsverarbeitung, Datenaufnahme, Datenlast und kurzlebige Daten erfüllen zu können.

Hohe Performance für festplattenbasierte Datenvolumen

SQL Server bietet flexible Auswahlmöglichkeiten für Tabellen- und Datenplatzierung mit In-Memory-Server- und flash-basierten Storage-Optionen. Dies ist von unschätzbarem Wert, da es der IT die Flexibilität bietet, Infrastrukturrressourcen bereitzustellen, welche die Leistungsanforderungen zu optimalen Gesamtbetriebskosten (TCO) erfüllen.

Storage-Pools und Daten-Tiering



Datenplatzierung und -konsolidierung

In einem Abschnitt weiter oben in diesem Whitepaper wurden die Datenvirtualisierung und die neuen Funktionen von SQL Server 2019 als einheitlicher Datenmanagementplattform für Daten in SQL Server und extern gespeicherte Daten erläutert. Die IT muss jedoch die Datenplatzierung für den gesamten Datenbestand im Blick haben.

Was den Storage betrifft, sollte die IT über eine Plattformstrategie verfügen, die die unvermeidlichen Landeplätze für persistente Daten unterstützen kann. Dies umfasst die Speicherung von Daten auf dem richtigen Storage-Typ, z. B. Scale-up, Scale-out, Datei-, Block-, Objekt- und langfristige Schutz- und Aufbewahrungsstrategien.

Eine weitere Überlegung betrifft die Datenkonsolidierung. Die überwiegende Mehrzahl der Microsoft SQL Server-Umgebungen umfasst viele Versionen, die über unterschiedliche Merkmale und Funktionen verfügen. Die Konsolidierung dieser Versionen in einem modernen All-Flash-Array ermöglicht die Auslagerung einiger Aufgaben an die Infrastruktur und sorgt für eine konsistentere und bessere Erfahrung in allen Versionen. Die erzielten Vorteile basieren größtenteils auf den vom Array bereitgestellten Datendiensten, die sich in der gesamten Branche in Bezug auf die Effektivität der Datendienste stark voneinander unterscheiden.

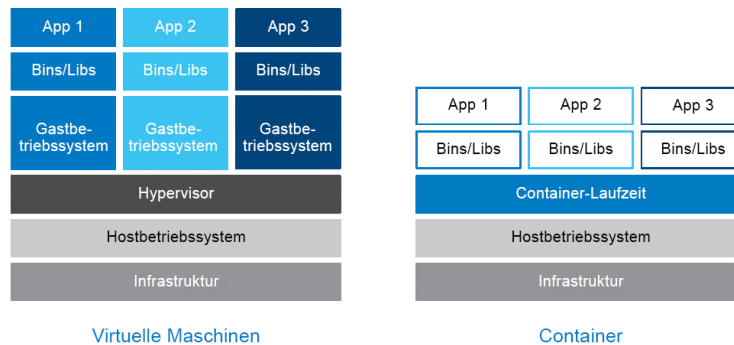
⁷ CTT, „[SATA vs SAS vs NVMe Performance Breakdown](#)“, Oktober 2018.

Im gesamten Dell EMC Produktportfolio gibt es verschiedene wichtige Alleinstellungsmerkmale und Unterschiede bei den Datendiensten.

- **Nutzbare Kapazität:** Ein häufig übersehener Aspekt ist der Overhead, den das System und das zugrunde liegende RAID verursachen oder nicht verursachen. Dell EMC-Systeme bieten einige der höchsten nutzbaren Kapazitäten in der Branche, wobei einige Systeme eine nutzbare Kapazität von bis zu 88 % bereitstellen.
- **Thin Provisioning:** Dies ist heute eine der wichtigsten Funktionen für die Maximierung der Kosteneffizienz einer Shared-Storage-Umgebung. Datenbankadministratoren (DBA) können so Datenbankdateien der zukünftig erwarteten Größe zuweisen und müssen die Größe später nicht erhöhen. Der Speicherplatz wird erst genutzt, wenn die Daten in das Array geschrieben werden.
- **Komprimierung:** Auch wenn dies im Dell EMC Portfolio an Storage-Arrays auf unterschiedliche Weise durchgeführt wird, bietet jedes System in der Regel eine bessere Komprimierungsrate als 2:1 für eine Datenbank, die in Microsoft SQL Server nicht bereits die Zeilen- oder Seitenkomprimierung nutzt.
- **Deduplizierung:** Es gibt viele Gründe, warum ein DBA oder Anwendungsbesitzer Datenbankkopien erstellen muss. Test/Entwicklung, Patch-Tests, nahezu Echtzeitanalysen oder einfach nur eine Kopie zum Schutz dieser Datenbankkopien werden im Array zu 100 % dedupliziert. Nur Änderungen, die an den Kopien oder der Produktion vorgenommen werden, führen dazu, dass neue Daten in das Array geschrieben werden.
- **Intelligente Snapshots:** So effizient wie die Deduplizierung auch ist, eine noch effizientere und schnellere Möglichkeit, eine Kopie zu erstellen, besteht darin, einfach einen Snapshot eines Volumes oder einer Gruppe von Volumes zu erstellen und die Snapshots dann auf einem anderen SQL-Server bereitzustellen. Durch die Erstellung von Datenbankkopien wird die Produktionsdatenbank nur wenig beeinträchtigt, wenn überhaupt, und es können selbst Kopien von sehr großen Datenbanken in weniger als fünf Minuten bereitgestellt werden.

Storage-Datendienste für Container

Container haben die Art und Weise, wie Unternehmen über die Entwicklung, Bereitstellung und Wartung von Anwendungen und Software denken, neu gestaltet. Relationale Datenbanken, einschließlich SQL Server, haben die Docker-Containerbildung zusammen mit dem Kubernetes-Orchestrierungs-Framework übernommen.



Container waren ursprünglich so konzipiert, dass sie von Natur aus kurzlebig oder statuslos sind, wodurch sie sich ideal für Test- und Entwicklungszwecke eignen. Container können jedoch auch für SQL Server-Datenbankumgebungen in der Produktion nützlich sein. Kubernetes und andere Container-Orchestratoren dehnen ihre Reichweite auf lang laufende Prozesse aus, insbesondere im Bereich Big Data und Analysen. Laut Gartner werden bis zum Jahr 2022 mehr als 75 % der globalen Unternehmen Container-Anwendungen in der Produktionsumgebung ausführen.⁸

Anwendungsentwickler arbeiten häufig außerhalb der Serverumgebungen, in denen ihre Programme ausgeführt werden müssen. Um Konflikte bei Bibliotheksversionen, Abhängigkeiten und Konfigurationseinstellungen zu minimieren, muss die Produktionsumgebung für die Entwicklung, Tests und die Integration vor der Produktion mehrmals neu erstellt werden.

⁸ Gartner, „6 Best Practices for Creating a Container Platform Strategy“. April 2019.

Das Dilemma beim Erstellen oder Aktualisieren von Datenbanken wie SQL Server besteht darin, dass die Daten persistent sein müssen und den Neustart, die Neuplanung oder das Löschen eines Containers überleben müssen. Wenn Container neu geplant werden, sollte der Storage ebenfalls verschoben und auf einem neuen Host zur Verfügung gestellt werden, damit der Container problemlos starten kann.

Um die Herausforderungen von statuslosen Containern und den Bedarf an persistenter Storage- und Replikationssoftware effektiv bewältigen zu können, unterstützt Dell EMC ein CSI-Plug-in (Containerized Storage Interface) für Kubernetes. Diese Integration ermöglicht Entwicklern, DBA und Storage-Administratoren die Nutzung der Schnittstellen, die sie am besten kennen, z. B. Befehlszeilenschnittstelle (CLI) und Kubernetes-Dashboard, zum Bereitstellen und Managen von an Dell EMC Storage-Datendiensten ausgerichteten persistentem Storage Tiering.

Die Akzeptanz von Containern und die Vorteile, die diese speziell für SQL Server bieten, werden weiter wachsen. Unternehmen, die Container-Datenbank-Workloads nutzen, können die Datenbankentwicklungszyklen sowie Produktionsbereitstellungen, Upgrades und die Orchestrierung hoher Verfügbarkeit erheblich vereinfachen. Ein wichtiger Pfeiler von skalierbaren Container-Cluster-Workloads ist die Big-Data-Cluster-Funktion von SQL Server 2019.

Netzwerkebene

Die heutigen Anwendungen und Datenbanken erfordern neue Denkweisen in Bezug auf die Netzwerkarchitektur. Die zugrunde liegende Netzwerkleistung und -auslastung kann sich negativ auf die Datenbankbeschleunigung und -skalierbarkeit auswirken. Beispielsweise kann der Anwendungsdurchsatz reduziert werden, wenn die Netzwerkbandbreite begrenzt wird.

Zur Unterstützung eines Datenbestands, der durch Datenlast und virtualisierte Daten gekennzeichnet ist, müssen IT-Abteilungen sicherstellen, dass Netzwerkplattformen und -topologien auf die zunehmende Menge, Vielfalt und Geschwindigkeit der Datenverschiebung vom Edge zum Core und zur Cloud und zurück vorbereitet sind.

Open-Networking-Technologien

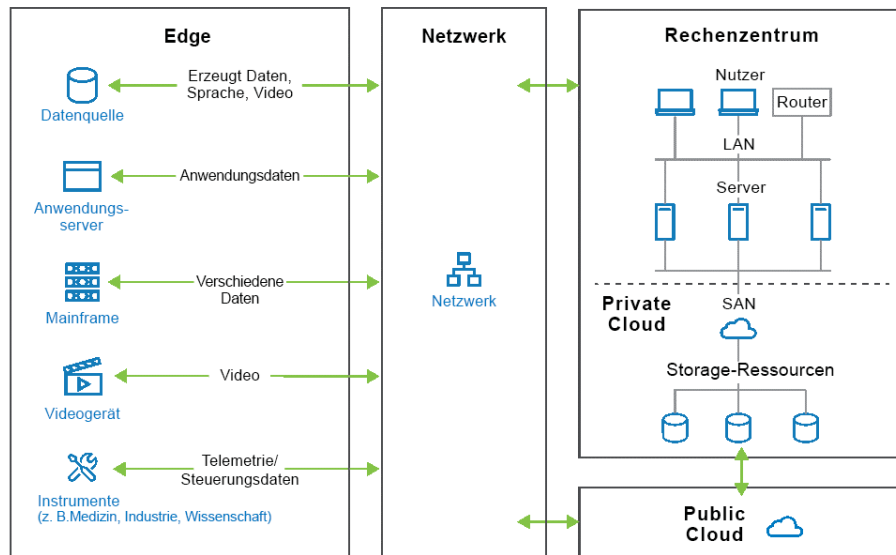
Bei der Vorbereitung der Unterstützung eines Datenbestands, der durch Datenlast und virtualisierte Daten gekennzeichnet ist, müssen IT-Abteilungen sicherstellen, dass Netzwerkplattformen und -topologien auf die zunehmende Menge, Vielfalt und Geschwindigkeit der Datenverlagerung vom Edge zum Core und zur Cloud und zurück vorbereitet wiederholen sind.

Open-Networking-Technologien, z. B. Software Defined Networking (SDN), ermöglichen es Unternehmen, ein Netzwerk bereitzustellen, das programmierbarer, flexibler und automatisierter ist, was wiederum folgende Vorteile bietet:

- Beseitigung von Netzwerkengpässen
- Vereinfachung der Einführung eines Cloud-Service-Bereitstellungsmodells
- Beschleunigung der Bereitstellung von Anwendungen
- Verbesserung der Recovery-Zeiten virtueller Maschinen (VM) zur Minimierung von Anwendungsausfallzeiten
- Entlastung der IT-Mitarbeiter, die sich dadurch stärker auf strategische IT-Weiterentwicklungen konzentrieren können
- Höhere Skalierbarkeit, sodass das Netzwerk die Anforderungen von Workloads, die wachsen und im Laufe der Zeit schwanken, bei Bedarf erfüllen kann

Laut Enterprise Strategy Group (ESG) müssen „Organisationen, die sich auf IoT-Initiativen (Internet der Dinge) vorbereiten und eine Verbindung zur Cloud herstellen, einen softwarebasierten Ansatz für die Netzwerktechnologie untersuchen. SDN trägt dazu bei, die Anforderungen an hohe Bandbreite und die der neuen Datenverkehrsmuster eines modernen digitalen Unternehmens zu erfüllen.“⁹

⁹ ESG, „[The Network's Foundational Role in IT Transformation](#)“, Mai 2018.



Darüber hinaus müssen mehrere Bereiche berücksichtigt werden, darunter:

Rechenzentrumsnetzwerk

Mit dem Aufkommen neuer Anwendungsfälle, die das Daten-Pipelining von IoT-Daten, Maschinendaten und Big Data umfassen, muss unbedingt sichergestellt werden, dass die Netzwerkverbindungen über ausreichend Bandbreite und Leistung für Spitzenlasten und die zunehmende Anzahl an Ost-West-Datenverkehrsmustern verfügen.

Storage-Netzwerk

Zur Leistungsoptimierung und Unterstützung von Business Continuity und Disaster Recovery haben große Unternehmen in der Regel FC- oder IP-basierte SAN-Storage-Systeme bereitgestellt, um gemeinsam genutzten Enterprise-Storage zu nutzen.

Da die Flash-Storage-Technologie weiter voranschreitet, nutzt die SQL Server-Umgebung die Vorteile der durchgängigen NVMe-Fabrics mit niedriger Latenz.

WAN (Wide Area Network)

Die Replikation ist zwar für WAN gegeben, bedenken Sie jedoch die zunehmende Rolle von WAN, wenn SQL Server-Umgebungen an verteilte Umgebungen angepasst werden. Das WAN wird eine wichtigere Rolle beim Sammeln von Daten aus Cloud- und Edge-Umgebungen spielen. Da immer mehr Unternehmen auf ein Hybrid-Cloud-Modell umsteigen möchten, benötigen sie eine Infrastruktur, die mit einer Umgebung verbunden ist, die einen Weg in die Cloud und Investitionsschutz für die nächste Hardware, die sie kaufen möchten, bietet. Wenn sie dann Workloads in die Cloud verschieben, ist die Verbindung bereits vorhanden. SD-WAN-Technologie (SDN in Wide Area Network) ermöglicht es Unternehmen, die Art und Weise zu modernisieren, wie Zweigstellen und Remotestandorte Verbindungen mit den von ihnen benötigten Daten herstellen. Diese Unternehmen können die Kosten drastisch senken, indem sie von teuren, separaten Leitungen (MPLS) auf kostengünstigere Breitbandkonnektivität oder eine Kombination aus mehreren Verbindungsoptionen umstellen.

Das Netzwerk, ob im Rechenzentrum oder in Access- und Edge-Netzwerken, ist ein kritischer Ausgangspunkt und die Grundlage der digitalen Transformation und Unterstützung der nächsten Entwicklung bei Datenmanagementlösungen wie SQL Server 2019. Die Netzwerkvision von Dell EMC¹⁰ basiert auf einem disaggregierten Modell mit einer offenen Umgebung, in der die Unternehmen aus einem breiten Angebot an innovativen branchenüblichen Netzwerkanwendungen, Netzwerkbetriebssystemen und Netzwerkhardwareprodukten auswählen können. Dieser Ansatz bietet den Kunden maximale Auswahl und Kontrolle über die Technologien, die sie auswählen, und die Architekturen, die sie übernehmen. Daraus resultieren messbare Kosteneinsparungen und eine höhere Serviceagilität.

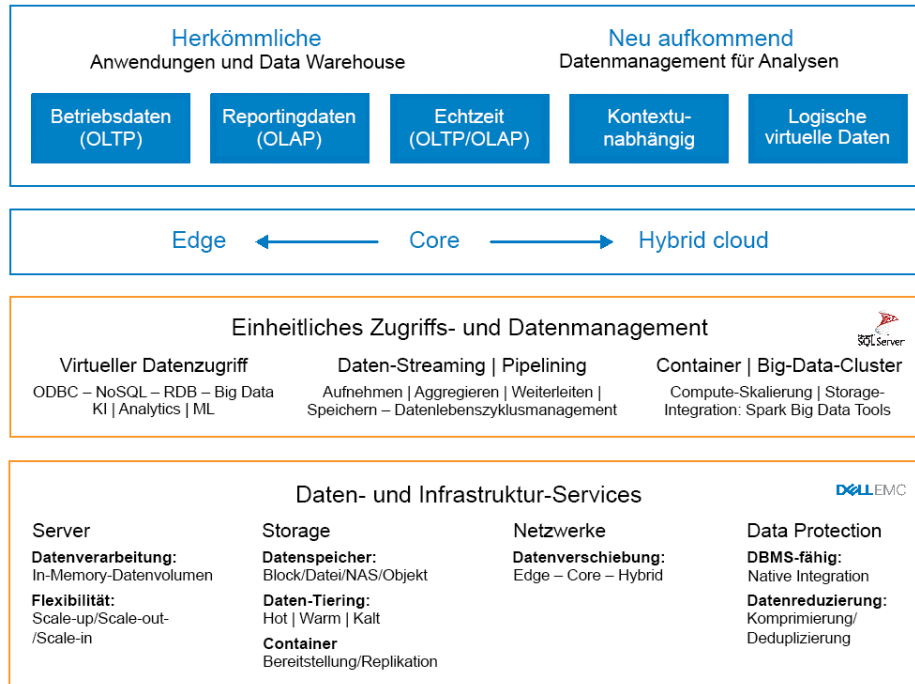
¹⁰ Weitere Informationen zu Dell EMC Open Networking finden Sie in der [Dell EMC Netzwerkübersicht](#).

Zusammenfassung

Nachdem wir nun ein Framework für neue Datenmanagementlösungen für virtualisierte Daten, wie im nachfolgender Abbildung dargestellt, betrachtet und Aspekte der Funktionen von SQL Server 2019 sowie Beispiele für ergänzende Datendienste und Infrastruktur-Anwendungsfälle, die durch Dell EMC ermöglicht wurden, betrachtet haben, werden im nächsten Abschnitt Überlegungen zur Planung der Umstellung vorgestellt.

Relationale Datenbankmanagementsysteme (RDMS)

Herkömmliche und neue Anwendungsfälle



Fazit

Wir befinden uns mitten in einem Paradigmenwechsel im Datenmanagement. Alle „IT-Wahrheiten“, die wir in den letzten Jahrzehnten erfahren haben, verändern sich – und zwar in eine positive Richtung. Die Art und Weise, wie wir gemeinsam Daten betrachten und was wir mit diesen Daten tun, wird in Zukunft zunehmend agiler und zugänglicher. Versierte IT-Verantwortliche wissen, dass dieses neue Zeitalter einen neuen Datenmanagementansatz erfordert. Einen Ansatz, der monolithische, isolierte Datenbanken durch neue Modelle für die Verwendung und gemeinsame Nutzung von Daten im gesamten Unternehmen ersetzt. Wenn Sie sicherstellen möchten, dass Ihre Planung aktiv auf diese Tendenzen eingeht, können Sie Entscheidungen treffen, die diesen Fortschritt weiter ermöglichen. Diese Entscheidungen werden zugunsten additiver und interaktiver statt eigenständiger isolierter Builds ausfallen, die zu Inseln vernachlässigter Technologie werden.

Die neue Wahrheit ist:

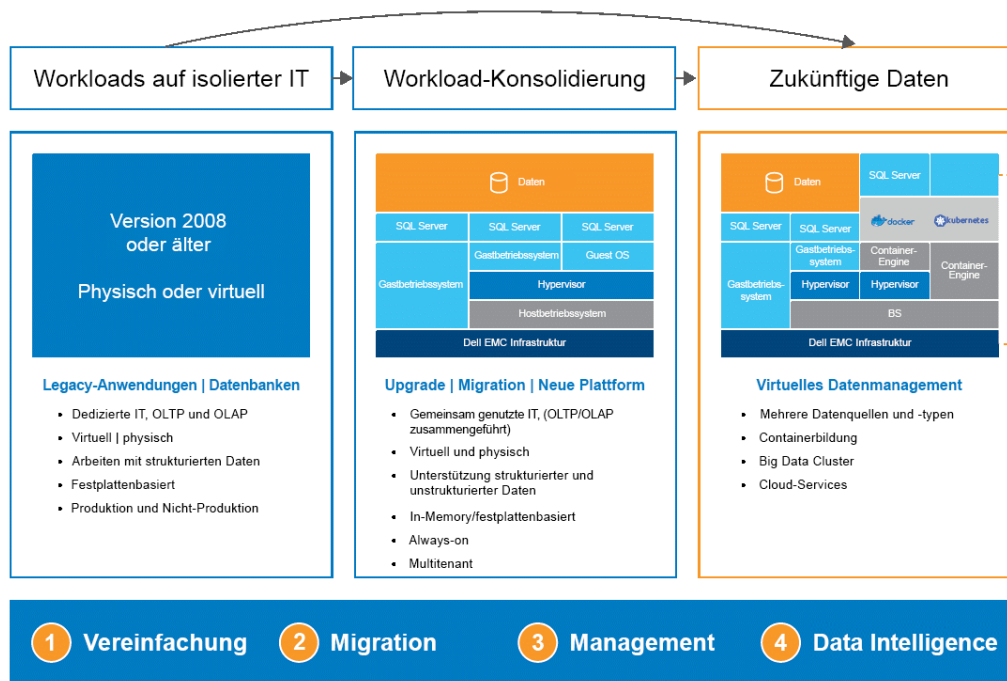
- Die Daten haben Schwerkraft („gravity“) und Anwendungen und Analysen kreisen um diese Daten, wo immer sie sich auch befinden mögen.
- Durch die Virtualisierung von Daten entsteht ein völlig neues Paradigma, wobei Daten demokratisiert und für alle verfügbar sind, ohne dass zeitaufwändige, komplexe ETL-Prozesse und immer größere Storage-Mengen benötigt werden.
- Die Mentalität des Komplettaustauschs von Systemen hat ein Ende. Es geht nicht mehr darum, sich für das Eine oder das Andere zu entscheiden, sondern um die Verbindung von beiden. Sie haben die Freiheit, das zu tun, was für Ihr Unternehmen sinnvoll ist.

Relationale Datenbanken, z. B. SQL Server 2019, wurden mit Architekturen und Tools entwickelt, mit denen Unternehmens- und Betriebsdaten mit KI, KL und DL zusammengebracht werden können. Diese Weiterentwicklung zu vereinheitlichten Datenmanagementplattformen ermöglicht es Unternehmen, mehr Daten aus mehr Quellen zu nutzen, sodass sie umsetzbare Erkenntnisse schneller und kostengünstiger als ihre Mitbewerber gewinnen können.

Wo startet Ihr Unternehmen?

Die digitale Transformation, die durch die IT-Transformation ermöglicht wird, ist eine Reise. Der Weg zum zukünftigen Datenbestand ist zwar einzigartig, wie unten dargestellt, doch die meisten haben einige gemeinsame Ziele. Dazu gehören die Aufrechterhaltung älterer Versionen von SQL Server-Bereitstellungen (einschließlich der Anwendungen) bei der Migration auf SQL Server 2019 und letztlich die Nutzung von Containern, Big Data-Clustern und Hybriden Cloud-Services für den virtuellen Datenzugriff und die Verwaltung.

Planung der Veränderung



Sehen wir uns einige dieser Ziele an:

- Das typische Unternehmen von heute führt Hunderte von Anwendungen aus und unterstützt mehrerer Datenbanken und gemischte Workloads. Konsolidierungsstrategien müssen berücksichtigt werden, um die Verfügbarkeit und Leistung geschäftskritischer Anwendungen sicherzustellen und gleichzeitig mit weniger Ressourcen niedrige Latenzzeiten zu gewährleisten. Die Einführung schnellerer, leistungsfähigerer CPUs und neuer Storage-Technologien ermöglicht es Unternehmen, Datenbanken zu konsolidieren, ohne die herkömmlichen damit verbundenen Risiken einzugehen. Die Konsolidierung und Vereinfachung gemischter Workloads, die in isolierten IT-Umgebungen mit Microsoft SQL ausgeführt werden, senkt nicht nur die Gesamtbetriebskosten (TCO) und steigert die Produktivität, sondern stellt auch eine moderne Infrastrukturgrundlage für zukünftige Daten bereit.
- Unternehmen möchten eine größere Auswahl bei Programmiersprachen, Datentypen und Betriebssystemen, um von ihrem Datenbestand zu profitieren und dessen Wert zu maximieren. Plattformumstellung von SQL Server auf Linux (oder eine gemischte Windows®-/Linux-Plattform) öffnet die Türen für mehr Flexibilität in Bezug auf Datenintegration, Containerbildung und Analyse. Es gibt auch potenzielle Kostenvorteile, einschließlich Lizenzen, Schulungszeit und Hardwarekosten.¹¹
- Die Containerbildung stellt eine flexible Plattform für selbst die komplexesten Anwendungen und Datenbanken dar. Die Containerbildung von SQL Server ermöglicht es Unternehmen, Aktualisierungen und Upgrades im Handumdrehen bereitzustellen, wobei sie die Flexibilität haben, sie lokal zu erstellen, in der Cloud bereitzustellen und überall auszuführen.

¹¹ Microsoft, „[Re-platforming and modernizing your data workloads with SQL Server on Linux](#)“, Juli 2018

- Entwickeln Sie eine Roadmap- und Bereitstellungsstrategie für den Wechsel zu Microsoft SQL Server 2019. Je nach Status des aktuellen Zustands und des gewünschten Zustands umfasst die Planung die Erstellung von Profilen und Bewertungen für Anwendungen und SQL Server-Datenspeicher zur Priorisierung von Projekten (Upgrades, anwendungsorientierte agile Migrationen und potenzielle Plattformänderungen), die in Phasen während der Umstellung ausgeführt werden.

Mit einer modernen Datenmanagementplattform und ergänzenden Daten- und Infrastrukturservices ist die IT bereit, das Unternehmen mit IT-Services für neue datengesteuerte intelligente Anwendungen, Geschäftsprozesse und Analysen zu unterstützen, die vom Edge zum Core ermöglicht werden.

Dell Technologies ist bereit, Sie auf dem Weg zu unterstützen

Microsoft und Dell Technologies sind seit über 30 Jahren Partner, um vorab getestete, vorab validierte und vollständig integrierte Lösungen bereitzustellen, die für eine optimierte Leistung für Microsoft SQL Server, Microsoft Azure Stack und Microsoft Azure Stack HCI entwickelt wurden. Microsoft gewährt Partnern Kompetenzen, die ihre Fähigkeiten sowohl durch qualifizierte Mitarbeiter als auch durch zufriedene Kunden unter Beweis stellen. Dell Technologies verfügt über mehr als 16.000 zertifizierte Mitarbeiter in den Bereichen Kunden-Support, Wartung, Bereitstellung, Schulung und Beratung und 18 globale Kompetenzen.

Dell Technologies besitzt das Fachwissen, um Sie bei der Umstellung auf virtualisierte Daten am optimalen Speicherort zu unterstützen, ohne dass Sie das gesamte System austauschen müssen. SQL Server verwaltet Ihre Daten auf verschiedenen Plattformen – On-Premise und in der Cloud. Das Ziel von Dell Technologies Consulting Services ist es, Sie dort abzuholen, wo Sie gerade stehen – auf jeder Plattform, mit den Tools und Sprachen Ihrer Wahl.

[Dell Technologies Consulting Services](#) unterstützen Sie bei der Planung eines soliden, grundlegenden Satz an Zielen und der Entwicklung einer Roadmap für die Modernisierung. Unsere SQL-Modernisierungsteams führen eine umfassende Bewertung durch, die Sondierungs-Workshops, Interviews und Vordenkerpositionen umfasst, um Ihnen Hilfestellung beim Erstellen der richtigen Datenmanagementstrategie zu bieten. Wir ermitteln Ihre Ziele im Hinblick auf die Zukunft und erstellen eine umsetzbare Roadmap, Nutzenanalyse und die anfängliche Migrationspriorisierung für Ihre wichtigsten Workloads. Wir unterstützen Ihr Unternehmen dabei, sein volles Potenzial mit einer Datenstrategie auszuschöpfen, die SQL Server, Azure SQL-Datenbank und Azure SQL Data Warehouse beinhalten kann.

Wir analysieren die aktuelle SQL Server-Umgebung, einschließlich des aktuellen Status aller zum Umfang gehörigen Server, der zugehörigen Workloads und Konfigurationen. Anschließend leisten wir Folgendes:

- Inventarisieren und Klassifizieren von Anwendungen, die an SQL Server-Datenbanken ausgerichtet sind, sowie aller Abhängigkeiten, Durchdenken von Verbindungen, Reporting, ETL-Prozessen und so weiter.
- Gruppieren und Priorisieren der SQL Server-Datenbanken oder -Instanzen nach Anwendungsgruppe, und Entwickeln eines kurzfristigen Modernisierungsplans und einer Roadmap für die Modernisierung. Dies ist auch ein ausgezeichneter Zeitpunkt, um die Datenbankkonsolidierung in Betracht zu ziehen.
- Ermitteln der groben Größenordnung zukünftiger Compute-, Storage- und Softwareanforderungen zur Unterstützung eines Modernisierungsplans.

Wenn Sie sich in einer Situation befinden, in der sich einige Ihrer Daten immer noch auf Datenbanken, deren Support-Zeitraum abgelaufen ist (z. B. SQL Server 2008), befinden und Sie kein Upgrade durchführen können, aber nicht riskieren möchten, keine weiteren Sicherheitsupdates mehr zu erhalten, migrieren Sie zu Azure oder Azure Stack, ohne Ihre Anwendungen zu verändern, wodurch Sie über weitere 3 Jahre Sicherheits-Updates erhalten.

Darüber hinaus können wir Ihre Daten zu Ihrem bevorzugten Ziel migrieren, weil wir verstehen, dass die sie flexibel genug sein müssen, um intelligente Anwendungen für allen Daten, auf jeder Plattform, in jeder Sprache On-Premise und in der Cloud zu erstellen. Wir unterstützen Sie auch bei der erfolgreichen Durchführung von Upgrades, Plattforummstellungen und/oder Konsolidierungen.

Weitere Informationen

Wenn Sie bereit sind, einen zertifizierten, preisgekrönten Microsoft-Partner in Betracht zu ziehen, der Ihre Modernisierungsbestrebungen mit SQL Server versteht, kann der ganzheitliche Ansatz von Dell EMC Ihnen dabei helfen, Risiken und Geschäftsunterbrechungen zu minimieren.

Weitere Informationen finden Sie unter www.dell EMC.com/SQL und erhalten Sie von Ihrem [Dell EMC Mitarbeiter](#).

Weiterführende Literatur

Dell EMC

- [Konsolidierung und Vereinfachung von gemischten Datenbank-Workloads mit Dell EMC](#)
- [Dell EMC Ready Solutions for Microsoft SQL: Design for Dell EMC XtremIO: With PowerEdge R840, Windows Server 2016, and RHEL 7.6](#)
- [SQL Server-Container unter Linux: Anwendungsbeispiele für die Softwareentwicklung mit Dell EMC Infrastruktur](#)
- dell EMC.com/sql

Microsoft

- [SQL Server on Linux: A guide to re-platforming and modernising your data workloads](#)
- [Microsoft SQL Server 2019 Big Data-Cluster](#)

Intel

- [Persistenter Intel Optane DC-Speicher](#)

Die Informationen in dieser Veröffentlichung werden ohne Gewähr zur Verfügung gestellt. Dell Inc. macht keine Zusicherungen und übernimmt keine Gewährleistung jedweder Art im Hinblick auf die in diesem Dokument enthaltenen Informationen und schließt insbesondere jedwede implizite Gewährleistung für die Handelsüblichkeit und die Eignung für einen bestimmten Zweck aus.

Copyright © 2019 Dell Inc. oder ihre Tochtergesellschaften. Alle Rechte vorbehalten. Dell, EMC und andere Marken sind Marken von Dell Inc. oder Tochtergesellschaften. Andere Marken sind möglicherweise Marken der jeweiligen Inhaber. Veröffentlicht in den USA. Referenznummer: H17964

Microsoft®, SQL Server®, Azure®, Excel® und Windows® sind entweder eingetragene Marken oder Marken der Microsoft Corporation in den USA und/oder anderen Ländern. Linux® ist eine eingetragene Marke von Linus Torvalds in den USA und anderen Ländern. Apache®, Spark® und Hadoop® sind Marken der Apache Software Foundation. Oracle® ist eine eingetragene Marke von Oracle und/oder deren Tochterunternehmen. Hortonworks® ist eine Marke von Hortonworks, Inc. in den USA und anderen Ländern. Cloudera® ist eine Marke oder eine Handelsaufmachung von Cloudera. Intel® und Optane™ sind Marken der Intel Corporation oder deren Tochtergesellschaften in den USA und/oder anderen Ländern. Linux® ist eine eingetragene Marke von Linus Torvalds in den USA und anderen Ländern. IBM® ist eine Marke oder eingetragene Marke der International Business Machines Corporation in den Vereinigten Staaten und/oder in anderen Ländern. MongoDB® und MongoDB® sind eingetragene Marken von MongoDB, Inc. SAP® und SAP HANA® sind eingetragene Marken der SAP SE in Deutschland und anderen Ländern.

Dell EMC ist der Ansicht, dass die Informationen in diesem Dokument ab dem Tag ihrer Veröffentlichung zutreffend sind. Diese Informationen können jederzeit ohne vorherige Ankündigung geändert werden.