

SQL Server 2019 – Neuerungen in der Übersicht

Mit der Veröffentlichung von SQL Server 2019 feiert Microsoft das 25-jährige Jubiläum des relationalen Datenbankenmanagementsystems (RDMS). In diesem Artikel möchten wir Ihnen die Änderungen der jüngsten Veröffentlichung erklären.

Das größte Augenmerk des Releases liegt vermutlich auf dem neuen Umgang bezüglich der Kompatibilität von SQL Servern mit verschiedenen Betriebssystemen und Programmiersprachen. Doch neben zahlreichen anderen Änderungen, hat auch die in SQL Server 2016 veröffentlichte **PolyBase** Neuerungen zu verzeichnen.

Im Bezug auf Performanz und Sicherheit macht Microsoft seine Vorteile deutlich: SQL Server 2019 wird sowohl im Bezug auf OLTP Performanz basierend auf TPC-Benchmarks, als auch im Bezug auf Data-Warehousing basierend auf TPC-H-Benchmarks als schnellstes RDMS bezeichnet.

Und auch an Sicherheit mangelt es nicht: Seit nun mehr 8 Jahren wird der SQL Server von Microsoft mit der geringsten Anzahl an gemeldeten Sicherheitslücken und Schwachstellen als das sicherste RDBMS im Vergleich zur Konkurrenz bezeichnet.

Aufgaben-kritische Performanz

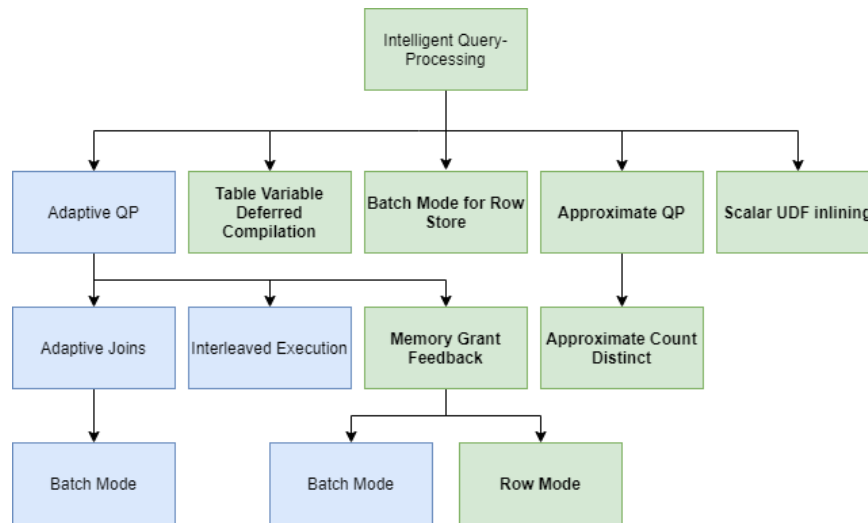
Persistenter Speicher

Mit der stetig wachsenden Anwendung von persistentem Speicher in der Industrie, ergeben sich viele Möglichkeiten alltägliche Aufgaben hinsichtlich der Performanz enorm zu verbessern. Mit dem SQL Server 2019 wird persistenter Speicher als Speicher-Medium in vollem Umfang unterstützt.

Intelligente Abfragen-Verarbeitung

Die **intelligente Abfragen-Verarbeitung**, deren Basis die Familie der zu betrachtenden Eigenschaften bildet, wurde verbessert. Die zur Verfügung stehenden Eigenschaften wurden erweitert und bieten damit eine Vielzahl an neuen Möglichkeiten für den **Query-Optimizer**, aus Abfrage-Ausführungen zu lernen. Damit verbessert sich beispielsweise die Performanz von Anwendungen, ohne eigenes Zutun.

In der folgende Graphik wird die erweiterte Familie der Eigenschaften dargestellt:



In-Memory TempDB

Die **TempDB** wird im SQL Server 2019 direkt im Speicher dargestellt. Daraus ergeben sich noch schnellere Verarbeitungszeiten für alle Aufgaben in ihrer Ausführung, auf die die **TempDB** zurückgreifen.

Verfügung über Daten

Daten-Virtualisierung

Mit der Verbesserung von **PolyBase** und **Daten-Virtualisierung** versucht Microsoft neuen Umgang mit verschiedenen Typen von Szenarien zu ermöglichen: Daten-Virtualisierung kann nun basierend auf unterschiedlichen Daten-Quellen realisiert werden. Dies sorgt dafür, dass alle benötigten Daten zur Ausführungszeit einer Abfrage zusammengeführt werden können, ohne eine **ETL-Pipeline** zu benötigen.

SQL Server Verteilung

Für die Verteilung des SQL Server hat sich Microsoft ein neues Muster einfallen lassen: Mit Hilfe von sogenannten "**Big Data Clusters**" lassen sich SQL Server Instanzen mit allen Fähigkeiten und Berechtigungen zusammen mit **Spark** und **HDFS** in einer integrierten Lösung ausliefern. Dies ermöglicht das besonders einfache Ausliefern von Umgebungen, in denen Abfragen auf großen Datensätzen getätigt werden (bis zu mehr als 100 Petabyte).

Es entsteht ein geteilter, skalierbarer **Data-Lake**, der sowohl von der SQL Server Instanz, als auch von **Spark** direkt angesprochen werden kann.

Vollständige KI/ML-Plattform

Der SQL Server 2019 bietet eine vollständige Plattform hinsichtlich Künstlicher Intelligenz im Bezug auf Maschinelles Lernen:

- × Daten können sowohl auf dem SQL Server oder im HDFS gespeichert werden.
- × Daten-Aufbereitung kann mit Hilfe von Spark oder von gespeicherten Prozeduren des SQL Servers realisiert werden.
- × ML-Model Training kann mit Hilfe der in Spark implementierten ML-Bibliothek, oder der Machine-Learning-Dienste der SQL Server Master-Instanz durchgeführt werden.

Zusätzlich besteht die Möglichkeit, ein ML-Modell als Repräsentation eines REST-API-Containers bereit zu stellen und diesen Container auf dem Daten-Cluster zu platzieren. Dies erleichtert beispielsweise Das Einreichen neuer Daten durch die Anwendungs-Entwickler.

Betriebssysteme und Container

Multi-Plattform

Seit dem Jahr 2017 wird Linux als Betriebssystem für den SQL Server unterstützt. Mit der Veröffentlichung des SQL Server 2019 stehen nun zusätzlich die folgenden Eigenschaften unter Linux zur Verfügung:

- × PolyBase,
- × Machine-Learning Dienste,
- × Distributed Transaction Coordinator (DTC),
- × Replikation

Mit diesen jüngsten Änderungen ist nun eine nahezu hundertprozentige Kompatibilität zwischen den Systemen auf Windows und Linux zu verzeichnen.

RHEL-Container

In Kooperation mit **redhat** stehen nun RHEL-basierte Container-Images zur Verfügung. Diese sind in der Container-Registry von Microsoft zu finden.

Always On – Kubernetes

Die **“Always On“**-Verfügbarkeitsgruppen lassen sich mit dem SQL Server 2019 direkt in **Kubernetes** darstellen und bieten damit entscheidende Vorteile bezüglich der skalierenden Lesevorgänge oder der Hoch-Verfügbarkeit.

Sicherheit und Vorschriften

Data Classification Engine

Mit SQL Server 2019 steht jedem Benutzer Data-Klassifizierung direkt **“out-of-the-box“** zur Verfügung. Dies ist vor allem bezüglich der neuen Regulationen hinsichtlich Kunden-Daten sinnvoll: Mit der Spezifikation einer Datenbank wird von der benutzten **Data-Classification-Engine** automatisiert alle unterschiedlichen Typen von Daten in der Datenbank erfasst und anhand der Ergebnisse ein Bericht erstellt.

Bei den unterstützten Daten-Typen handelt es sich um beispielsweise um **PCI-** oder **GDPR-Daten**, es besteht jedoch zusätzlich die Möglichkeit, eine benutzerdefinierte Klasse von Regeln anzulegen und diese zu verwenden.

Always Encrypted

Hinsichtlich der **“Always encrypted“** Funktionalität besteht nun die Möglichkeit auf Client-seitige Verschlüsselung zurückzugreifen. Dies ermöglicht die saubere Trennung zwischen Datenbanken-Administratoren und Anwendungs-Entwicklern/Benutzer. Die Daten können nicht innerhalb der Datenbank entschlüsselt werden.

Sichere Enklaven

Eine weitere Neuerung im Bezug auf die **“Always encrypted“** Funktionalität ist, dass die Verschlüsselung der Daten nun in abgesicherten **Enklaven** stattfinden kann.

Weitere Informationen bezüglich des Umgang mit Enklaven sind [hier](#) zu finden.

Entwickler und DBA Werkzeuge

Notebooks

Im Rahmen des **"Azure Data Studio"** stehen nun sog. **Notebooks** zur Verfügung. Notebooks sind Dokumente in denen sich eine Komposition von Text und Code besonders leicht realisieren lässt. Der große Vorteil ist, dass der enthaltene Code direkt im Dokument ausgeführt werden kann. Dies ermöglichen das Erstellen von "interaktiven" Dokumenten. Sie eignen sich besonders gut für Kollaboration oder zum Dokumentieren von Standardverfahren zur Fehleranalyse oder ähnlichem. Mit der zusätzlich zur Verfügung stehenden Integration von **Git**, lassen sich Notebooks sogar kollaborativ teilen.

Datenvisualisierung (SandDance)

Eine weitere Neuerung des **"Azure Data Studio"** ist die Integration von **SandDance**, einem Service, der ursprünglich im **Microsoft Research** Programm zur Verfügung stand. **SandDance** bietet die Möglichkeit, innerhalb des Azure Data Studio **ad hoc Visualisierungen** von Daten zu erzeugen.

Dieses neue Feature, sowie die zuvor enthaltenen Fähigkeiten des Azure Data Studios stehen sowohl für SQL Server Instanzen als auch für **PostgreSQL** und **MySQL** zur Verfügung.