

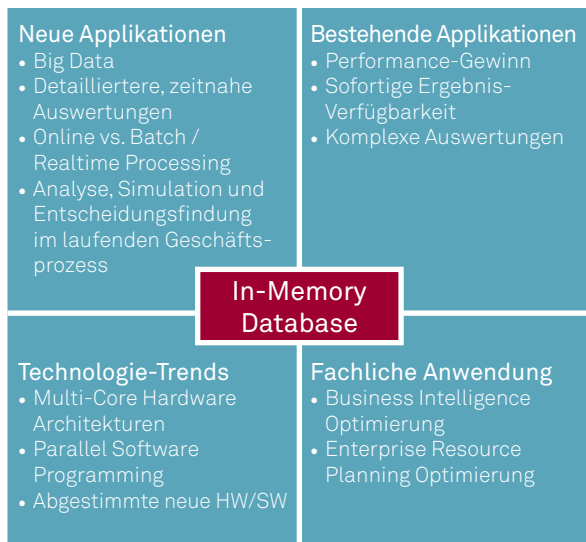
In-Memory Datenbanken

Next Generation Database Technology

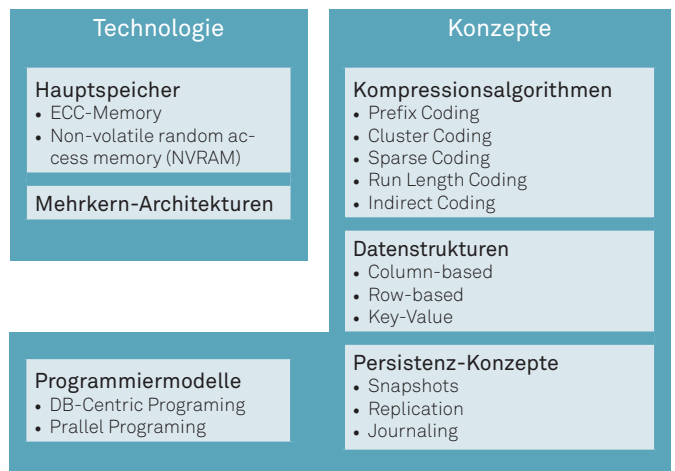
In-Memory-Datenbanken verarbeiten Daten ausschließlich im Hauptspeicher und erreichen so signifikant höhere Geschwindigkeiten als herkömmliche Datenbank-Systeme, die mit Festplatten arbeiten. Durch die Möglichkeit, immer größeren Hauptspeicher in Verbindung mit einer Komprimierung zu nutzen, können Datenvolumina bis zu mehreren Terabytes verwaltet werden. Der Markteintritt von SAP mit der eigenen HANA-Technologie beeinflusst direkt den Umgang mit den Applikationen in allen Branchen durch neue Anwendungen.

Definition

In-Memory-Datenbanken beschleunigen den Zugriff auf Anwendungsdaten im Vergleich zu traditionellen Datenbanken erheblich, da sämtliche Daten während der Verarbeitung im Hauptspeicher gehalten werden. Das liegt zum einen am schnelleren Zugriff auf den Hauptspeicher im Vergleich zu herkömmlichen Festplatten. Zum anderen ergänzen die Trends hin zu Mehrkern-Architekturen und Parallelprogrammierung den Ansatz der In-Memory Datenverarbeitung.



Die Kombinationen dieser Ansätze ermöglicht einen hohen Performancegewinn beim Lesen und Schreiben der Daten, vor allem aber bei Aggregationen über den gesamten Datenbestand.

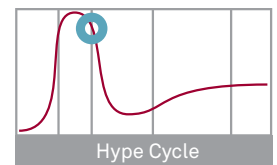


Mittels Komprimierung und einer spaltenbasierten Abfragestruktur, die keinen Speicherplatz für leere Felder beansprucht, ist es zudem möglich, Datenbestände im Terabyte-Bereich hochperformant zu verarbeiten.

Um die bei Datenbanken oftmals gewünschten ACID-Eigenschaften – Atomicity, Consistency, Integrity and Durability – für Anwendungen gewährleisten zu können, werden zusätzlich Festplatten oder spezielle Speichermodule (NVRAM) für die Persistenz eingesetzt.

Reifegrad

Die Ausnutzung des Hauptspeichers, um die Performance zu erhöhen, wird seit Jahren im Datenbank-Umfeld durch verschiedenste Caching-Me-



chanismen erfolgreich genutzt. Obwohl das erste kommerzielle Produkt 1993 auf den Markt gekommen ist, sind In-Memory Datenbanken bis dato aber eher sehr spezialisiert eingesetzt worden und müssen sich im breiten Anwendungsfeld noch beweisen.

Da viele NoSQL-Produkte auch den In-Memory-Ansatz verfolgen, wird die Verbreitung in den nächsten Jahren deutlich zunehmen.

Marktübersicht

Alle großen Datenbankanbieter haben mittlerweile In-Memory Technologien im Produktportfolio. Prominenteste Vertreter sind hier TimesTen von ORACLE, SolidDB von IBM und die HANA Appliance von SAP. Zudem existieren viele Open-Source-Produkte auf dem Markt, wie beispielsweise H2 und Redis.



Der In-Memory Ansatz ist somit im Markt platziert und spielt seit vielen Jahren eine feste Rolle im Business-Analytics-Umfeld. Hier ist beispielsweise das Produkt QlikView der Firma QlikTech seit über 10 Jahren erfolgreich am Markt.

Alternativen

Alternativ können traditionelle Datenbanken mit massivem Caching im Hauptspeicher und Persistenz auf Solid-State Disks (SSD) zum Einsatz kommen. Diese Optimierungen reichen meist aber nicht an die Performance von In-Memory Ansätzen heran und benötigen deutlich mehr

Speicherplatz als einige dedizierte In-Memory Datenbanken mit Ihren Kompressionsalgorithmen.

Referenzszenario

Die Einsatzgebiete für In-Memory Ansätze liegen zum einen in der Beschleunigung bestehender Applikationen. Hier seien vor allem Reporting und Auswertungen (Business Intelligence) genannt. Zum anderen werden vorher nicht erreichbare, detailliertere Analysen und die Umsetzung von Realtime-Auswertungen – selbst bei größeren Datenmengen – durch die neue Technologie ermöglicht.

Mittlerweile werden aber nicht nur die analytischen Daten für Auswertungen betrachtet, sondern auch die operativen Daten, was die Reaktionsfähigkeit der Unternehmen erhöht (Real-Time Enterprise). Ein Beispiel dafür ist der Einsatz von SAP HANA: HANA ist heute die Basis für die SAP Business-Intelligence-Lösung und zukünftig auch für die SAP-ERP-Lösung.

Business Impact

Durch In-Memory Datenbanken entstehen neue Anwendungen und fachliche Anwendungsfälle, die durch die Performance der Systeme erst möglich werden. Die Realtime-Auswertungen tragen zur schnelleren Entscheidungsfindung in den Prozessen der Unternehmen bei.

Pro	Contra
Extreme Lese- und Schreib-Performance	Verzicht bzw. Mehraufwand für gewohnte Garantien der RDBMS – speziell der ACID-Eigenschaft „Durability“
Ausnutzung des Trends hin zur Mehrkern-Architektur	Je nach Lösung (zum Beispiel SAP HANA) spezielle Hardwarekonfiguration nötig
Ausnutzung des Trends hin zur Parallelprogrammierung	Je nach Hersteller sind Anpassungen in den Anwendungen nötig
Möglichkeit des performanten Umgangs mit Big Data (jedoch nur im Terabyte Bereich)	Big Data im Petabytebereich ist aktuell als In-Memory noch nicht möglich.
	Verlust der ursächlichen Vorteile bei Clustering.

msg systems ag

Robert-Bürkle-Straße 1 | 85737 Ismaning/München
 Telefon: +49 89 96101-0 | Fax: +49 89 96101-1113
 www.msg-systems.com | info@msg-systems.com

Stand: September 2012

<http://www.msg-systems.com/techrefresh>

