In-Memory-Datenbanken

Eine **In-Memory-Datenbank (IMDB)** (auch **speicherresidente Datenbank** oder **Hauptspeicherdatenbank**, im Englischen auch **Main Memory Database**) unterscheidet sich von anderen Datenbanken insofern, als dass sie ihren Datenbestand primär im RAM hält.

Die Motivation hinter dem Konzept der In-Memory-Datenhaltung ist klar. Eine Datenbank, bei der die Daten im Hauptspeicher gehalten werden, bietet einen erheblich schnelleren Datenzugriff als herkömmliche Systeme, bei der die Daten auf magnetischen Festplatten gespeichert sind. Die Manipulation der Daten findet hier direkt im Hauptspeicher statt, ohne dass sie zuvor von der Festplatte geladen werden müssen. Es entfällt also ein langwieriger Zugriff mittels mechanischer Elemente zum Suchen, Speichern oder Laden der passenden Sektoren, der bis zu mehreren Millisekunden dauern kann. Der Zugriff auf RAM ist bis um den Faktor eine Millionen Mal schneller als bei Festplatten. Das ermöglicht eine extrem schnelle Verarbeitung, wo einige Auswertungen früher Stunden oder Tage gedauert haben.

„Reine“ Hauptspeicherdatenbanken sind darauf ausgelegt, keinerlei Plattenzugriffe zu benutzen, da sie dafür entwickelt wurden, den kompletten Datenbestand im RAM zu verwalten. Diese Tatsache macht sie zu etwas Besonderem, da sie so im Vergleich zu einer herkömmlichen Datenbank, die auf einer RAM-Disk implementiert wurde, immer noch schneller arbeitet, eben, weil sie im Gegensatz zu anderen Datenbankensystemen keinerlei Befehle für Plattenzugriffe oder aufwendige Cachingverfahren verwendet (was sie ja immer noch tun, auch wenn sie auf einer RAM-Disk operieren).

Klassische Festplatten sind eine der wenigen Einheiten in modernen Rechnern, in denen noch immer mechanische Bauteile verwendet werden. Das macht sie besonders langsam und so sind sie oft der Flaschenhals eines ansonsten modernen Systems. Der Trend geht dahin, derartige Speichermodule nicht mehr zur aktiven Datenhaltung zu verwenden, sondern als Backupmedien zu nutzen. Frei nach dem Datenbank-Experten Jim Gray lautet das Motto: „*Tape is Dead Disk is Tape Flash is Disk RAM Locality is King*“ (Gray 2006) (zu Deutsch etwa: „*Das Bandlaufwerk ist tot, die Festplatte ist das (neue) Bandlaufwerk, Flash ist die (neue) Festplatte, RAM ist das einzig Wahre*“).

Verwendung

Verwendung finden IMDBs in Bereichen, bei denen eine zeitnahe Datenauswertung erfolgen muss und Echtzeitanalysen gefragt sind oder auch in solchen Systemen, die über keinerlei persistenten Speicher verfügen (Router, Switches, etc.). Sehr oft werden In-Memory-Technologien auch in Kombination mit NoSQL-Technologien verwendet, wie etwa bei spaltenorientierten Datenbanken, wodurch die Anwendungen zusätzlich beschleunigt werden können. Im Zusammenhang mit Key-Value-Stores ist auch ein Einsatz als Cache-Speichersysteme für andere Datenbanken möglich.

Nachteile

Preis

Wenn RAM nun aber den Festplattenspeicher ersetzt, bedeutet das zwar einerseits einen erheblichen Geschwindigkeitsvorteil, andererseits jedoch auch höhere Hardwarekosten. Trotz steigender Kapazitäten und sinkender Preisen in den letzten Jahrzehnten ist Hauptspeicher immer noch wesentlich teurer als Festplattenspeicher. Eine Alternative bieten hier Flash-Speichermedien. Die Zugriffszeiten bei SSDs bewegen sich zwar im Mikrosekundenbereich, sind aber noch immer nicht so schnell wie Hauptspeicherzugriffe und auch hier ist das Preis-Leistungs-Verhältnis noch nicht so optimal wie beim Festplattenspeicher.

Persistenz

Doch nicht nur die Kosten sind ein Nachteil von IMDBs. Da es sich bei RAM um einen flüchtigen Speicher handelt, der seinen Inhalt beim Herunterfahren oder Abstürzen verliert, sind Sicherheitsmaßnahmen zu treffen, um einen persistente Datenbestand gewährleisten zu können. Um dies zu erreichen bieten sich ein paar Möglichkeiten. Darunter u.a.:

- Regelmäßige Erstellen von Snapshots auf nichtflüchtigem Speicher
- Zusätzliches Führen eines Logfiles, das alle Transaktionen protokolliert (mit den Snapshots zur Wiederherstellung des letzten persistenten Zustands)
- Erstellen von Replikationen, die bei einem Ausfall einspringen
- NVRAM mit eigener (unabhängiger) Stromversorgung

(Vgl. McObject 2014).

From:
https://wi-wiki.de/ - **Wirtschaftsinformatik Wiki - Kewee**

Permanent link:
https://wi-wiki.de/doku.php?id=bigdata:inmemory

Last update: **2015/10/05 21:14**