ChannelPartner

Link: https://www.channelpartner.de/a/aus-big-data-wird-small-data,2539460

DB2 mit BLU-Beschleunigung

Aus Big Data wird Small Data

Datum: 03.06.2013
Autor(en): Klaus Manhart

Mit BLU bietet IBMs neues Datenbanksystem DB2 eine ganze Reihe von Beschleunigungstechnologien für große Datenmengen. Mehrfaches, intelligentes Sortieren und Komprimieren reduziert Terabyte-große Daten auf ein Minimum. Davon profitieren besonders massiv lesende Zugriffe und typische Data-Warehouse-Anwendungen.

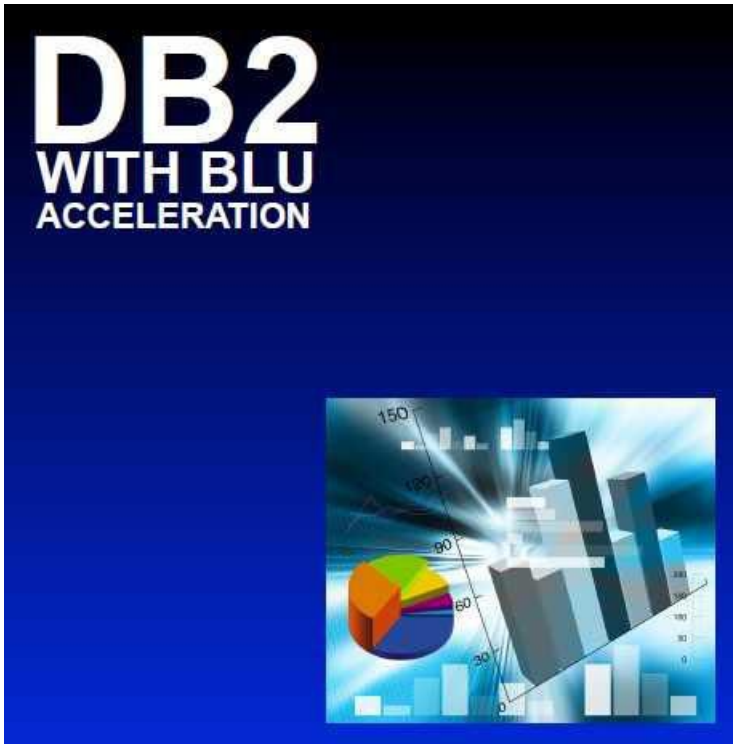


Foto: IBM

Große Datenmengen brauchen neue Technologien. An Terabyte und Petabyte von Daten versagen herkömmliche Methoden. In der neuen Version des Datenbankservers DB2 10.5 mit BLU Acceleration setzt IBM auf eine Technologie namens BLU, um auf Daten von analytischen Systemen In-Memory und damit sehr performant zuzugreifen.

BLU (Blink Ultimate), das auf das **Projekt Blink**[1] des IBM-Forschungszentrums in Almaden zurückgeht, kombiniert gleich mehrere Technologien: In-Memory-Computing, spaltenorientierte Datenbank-Tabellen, Data Skipping, SIMD (Single Instruction Multiple Data) und aus dem Supercomputing stammendes Vektorrechnen. Mit diesem wahren Feuerwerk an neuen Methoden soll die zu verarbeitende Datenmenge von mehreren Terabytes auf eine kleine Größe komprimiert - und damit die Datenverarbeitung und -analyse deutlich beschleunigt werden.

Reduziert werden die Daten dabei nicht, sondern lediglich sortiert und komprimiert. "Die eigentliche Datenmenge bleibt auf dem ursprünglichen Speichersystem erhalten", sagt Holm Landrock von der Experton Group in einer Review des neuen IBM-Systems. "Der Clou ist die mehrfache Datenkomprimierung auf dem Weg zum SQL-Befehl. Die Datenkomprimierung findet im Speicher und in der Analyse statt."

Vor allem zwei Grundziele kennzeichnen BLU:

- Die Daten sollen möglichst effizient im Memory gehalten werden.
- Die Daten werden in den Tabellen und im Speicher spaltenorientiert abgelegt.

BLU: Spalten statt Zeilen

DB2 BLU legt grundsätzlich die Daten spaltenorientiert im RAM (Random Access Memory) ab. Damit ist es bei bestimmten Operationen deutlich schneller als die bisherigen zeilenorientierten DB2-Systeme.

Der technische Hintergrund: Eine zeilenorientierte Datenbank hängt alle Datenwerte einer Zeile aneinander, es folgt die nächste Zeile und so weiter, wie im folgenden Beispiel mit Personalnummer, Nachname, Vorname, Gehalt demonstriert:

1,Huber,Josef,30000;2,Müller,Maria,50000;3,Meier,Sabine,45000;

Eine spaltenorientierte Datenbank hingegen geht Spalte für Spalte vor:

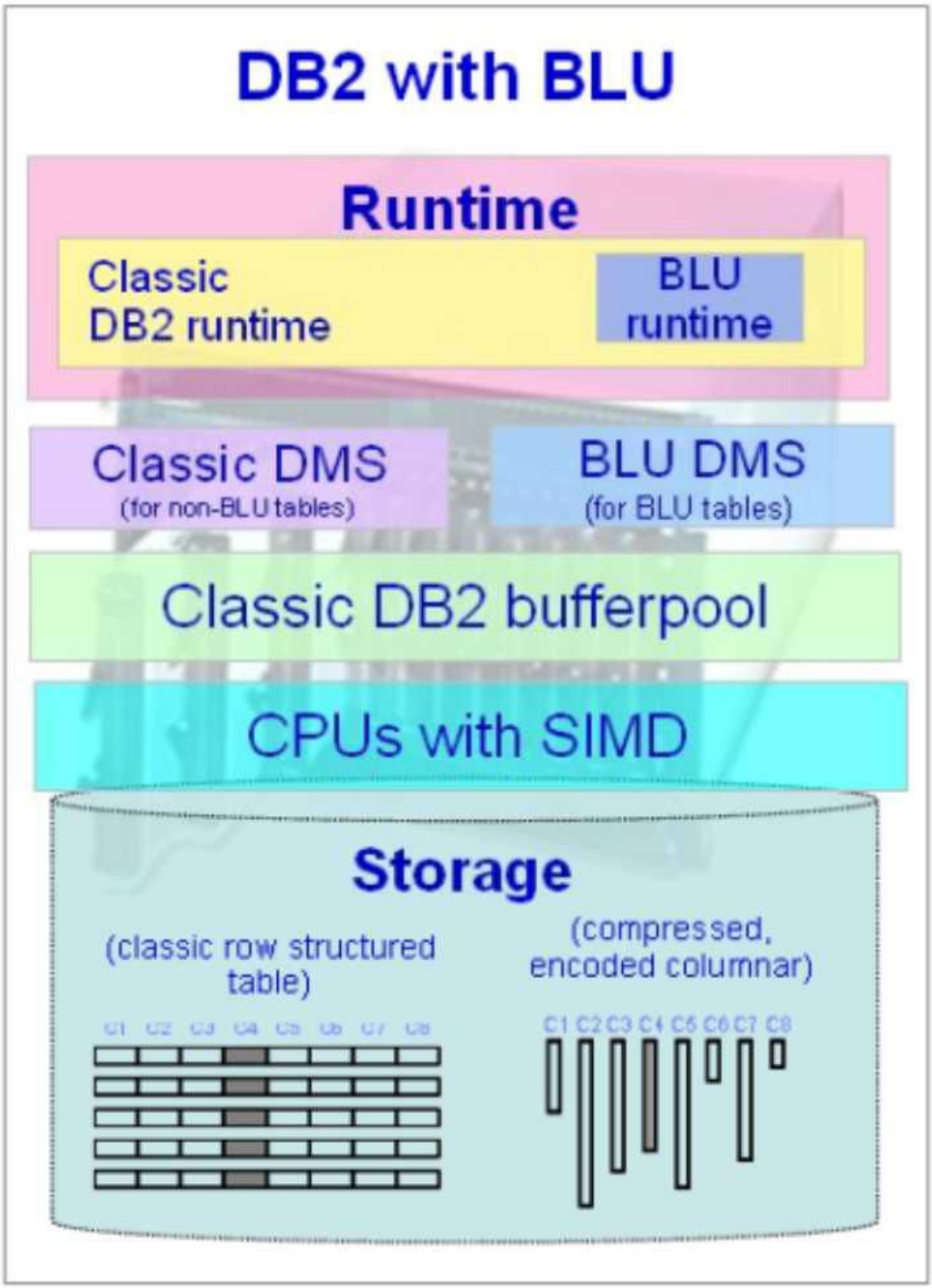
1,2,3;Huber,Müller,Meier;Josef,Maria,Sabine;30000,50000,45000;

Bei Abfragen in spaltenorientierten Systemen müssen nicht alle Zeilen, sondern nur die Spalten durchsucht werden, die für die Auswertung notwendig sind. Dieses nur scheinbar nebensächliche Implementierungsdetail hat zur Folge, dass sich gewisse Arten von Abfragen damit effizienter abarbeiten lassen. Das sind im wesentlichen Aggregierungen von Werten über viele Zeilen hinweg, wie sie in Data Warehouses und anderen analytischen Applikationen häufig vorkommen.

Da Spaltendaten einen einheitlichen Typ haben, stehen in spaltenorientierten Systemen zudem Komprimierungsoptionen zur Verfügung, die bei zeilenorientierten Daten nicht möglich sind. Bei der Konvertierung in Spalten werden die Daten deshalb in BLU gleichzeitig per Bitverschlüsselung komprimiert.

Durch die spaltenorientierte Speicherung der Daten in den Tabellen wird eine zusätzliche Komprimierung um etwa Faktor 5-10 erzielt. Erste Erfahrungen bei Kunden haben laut IBM gezeigt, dass die Ausführungszeiten von langlaufenden SQL Queries um Faktor 25 - 40 verkürzt werden konnten.

Caches und Buffer



Smart Caching und SIMD werten Daten hardwarenah In Memory und speichereffizient aus.
Foto: IBM

Auch beim Storage arbeitet die neue Datenbankversion deutlich effizienter: Mittels Smart-Caching, Single-Instruction-Multiple-Data (SIMD), Data-Skipping und einiger weiterer Funktionalitäten erfolgt die Auswertung der Daten hardwarenah In Memory und speichereffizient (siehe Grafik).

Beim **Smart Caching** werden die Daten in der bitorientierten, komprimierten Form im L1-L2-Cache gehalten, dem Pufferspeicher also mit den schnellsten Zugriffszeiten. Von dort werden sie dann direkt von der CPU ausgelesen und verarbeitet. Durch diese Datenhaltung ganz nahe an der CPU müssen die Daten vom Prozessor nicht aus dem RAM gelesen werden, was Zeit spart.

Bei den neueren x86- und Power-CPUs unterstützt BLU die **SIMD-Technologie** (Single-Instruction-Multiple-Data). SIMD verarbeitet mit einem einzigen Maschinenbefehl mehrere Daten gleichzeitig in der CPU. Diese Abarbeitung einer Vielzahl von Daten mit einer einzelnen Instruktion erhöht die Performance von DB2 weiter.

Im neuen DB2 **Buffer-Pool**, der ebenfalls die spaltenorientierten Daten komprimiert vorhält, cacht schließlich ein hocheffizienter, neuer Algorithmus die Daten. Gecacht werden nicht die ganzen Tabellen, sondern nur diejenigen Spalten, die für die Applikationen relevant sind. Beim SELECT-Statement:

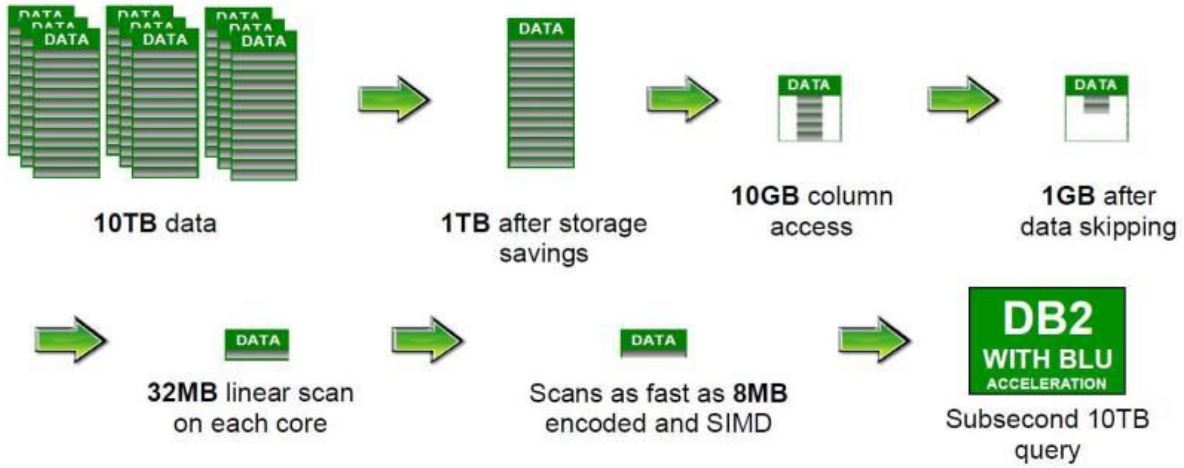
SELECT SPALTE1 FROM TABLE

werden also nur die Pages der 1. Spalte gepuffert und nicht der ganze Tabelleninhalt.

"Das ist ein sehr großer Unterschied im Vergleich zum Buffer-Pool-Zugriff bei den zeilenorientierten Tabellen", sagt Stefan Hummels, Senior IT Specialist und DB2-Experte bei IBM. "DB2 hat mit BLU eine sehr effiziente Speicherung der Daten im Cache und kann ihn deutlich besser nutzen. Das ist auch ein Grund dafür, dass DB2 BLU nicht so viel Hauptspeicher braucht, wie die Datenbank groß ist."

Ein Beispiel: Von 10 TB auf 8 MB

- **The system – 32 cores, 10TB table with 100 columns, 10 years of data**
- **The query:** `SELECT COUNT(*) from MYTABLE where YEAR = '2010'`
- **The optimistic result: sub second 10TB query! Each CPU core examines the equivalent of just 8MB of data**



Das Beispiel zeigt, wie BLU-Technologien 10 TB große Daten auf 8 MB reduzieren können.
Foto: IBM

Die vorgestellten BLU-Technologien greifen alle ineinander und ergänzen sich. Die Grafik zeigt anhand eines Beispiels, wie sich damit (riesige) 10 Terabyte Daten auf 8 MB reduzieren lassen.

Die 10 TB große Ausgangstabelle wird durch die erwähnten Komprimierungstechnologien um Faktor 10 reduziert, so dass sie nur noch eine Größe hat von

10 TB / 10 = 1 TB.

Über die Spaltenverarbeitung lässt sich die 1-TB-Tabelle weiter verkleinern. Besitzt die Tabelle beispielsweise 100 Spalten - im Data-Warehouse-Umfeld keine Seltenheit - und greift man nur auf eine dieser Spalten zu, haben sich die zu verarbeitenden Daten weiter verringert, nämlich von 1 TB auf

1 TB / 100 Spalten = 10 GB.

Für eine weitere Datenreduktion sorgt eine Technologie namens **Data Skipping**. Data Skipping legt automatisch in den Daten bei jedem 2000. Datensatz den Minimum- und Maximum-Wert der folgenden 2000 Datensätze ab. Will man beispielsweise Daten aus dem Jahr 2012 auswerten, dann springt DB2 via Data Skipping gleich zum Jahr 2012 und beginnt erst dort mit dem Lesen. Sind in der Datenbank 10 Jahre an Daten enthalten, dann werden von den 10 GB nur noch

10 GB / 10 =1 GB

an Daten gelesen.

Verfügt das Datenbanksystem über 32 Prozessorkerne, bearbeitet jeder Core nur noch

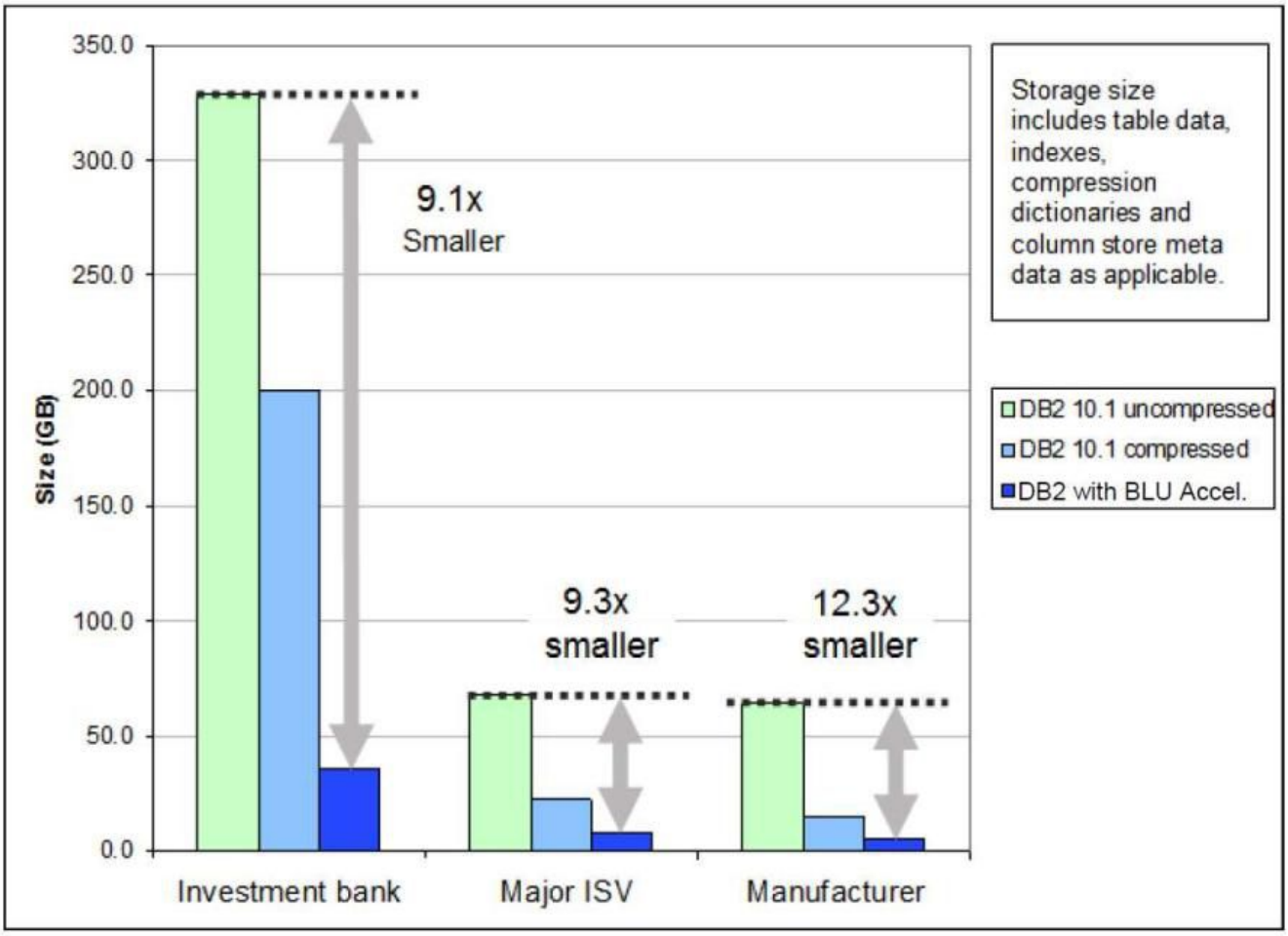
1 GB / 32 Cores = 32 MB.

Und schließlich: Mit der erwähnten SIMD-Technologie, die mit einer Instruktion gleichzeitig 4 Datenwerte bearbeitet, werden mit einem Maschinenbefehl dann nur noch

32 MB / 4 = 8 MB

verarbeitet. Von ursprünglich 10 TB Daten bleiben so am Ende nur noch 8 MB übrig. Damit wird das Datenbanksystem deutlich effizienter: Man braucht viel weniger als 10 TB an Hauptspeicher, und die Verarbeitung selbst wird performanter.

Load and Go



Die Grafik zeigt, um wie viel die Tabellen bei der Umwandlung von Zeilen in Spalten schrumpfen können.
Foto: IBM

Im Leistungsvergleich von DB2 10.5 BLU mit ihrem Vorgänger 10.1. zeigten sich bei ausgewählten Kunden deutliche Performance-Steigerungen in den Datenbank-Zugriffen. Die Verbesserungen bewegen sich nicht mehr nur im Prozentbereich, sondern im Bereich von Faktoren. Im Durchschnitt werden die Zugriffe um das 10- bis 25-fache beschleunigt, wobei IBM auch von Ausreißern mit Faktor 46 und 100 berichtet.

Die Grafik oben demonstriert, um wie viel die Tabellen bei der Umwandlung von Zeilen in Spalten schrumpfen. Dabei gilt grundsätzlich: Je größer die Tabelle, desto größer der Nutzen der Spaltenorientierung. Ein durchschnittlicher Komprimierungsfaktor von 10, wie es die Beispiele nahe legen, bedeutet, dass eine Tabelle, die ursprünglich 1 Terabyte groß war, auf 100 GB schrumpft. "Damit befindet man sich in ganz anderen Regionen und hat viel bessere Antwortzeiten", so Stefan Hummel.

BLU vereinfacht auch die Arbeit für den Datenbankadministrator. In klassischen Data-Warehouse-Systemen muss er festlegen, wie die Tabellen partitioniert und welche Daten komprimiert werden. Er muss Tabellen anlegen, Daten laden und in einem immer wiederkehrenden Prozess Indizes verwalten, die Datenbank tunen und Statistiken erstellen.

Da es in spaltenorientierten Tabellen keine Indizes gibt, entfällt in BLU das ganze Tuning und damit alles bis auf das Erzeugen von Tabellen und das Laden der Daten:

Database Design and Tuning:

• Decide on partition strategies

• Select Compression Strategy

• Create Auxiliary Performance Structures

• Materialized views

• ENTFÄLLT: Create indexes

• ENTFÄLLT: B+ indexes

• ENTFÄLLT: Bitmap indexes

• ENTFÄLLT: Tune memory

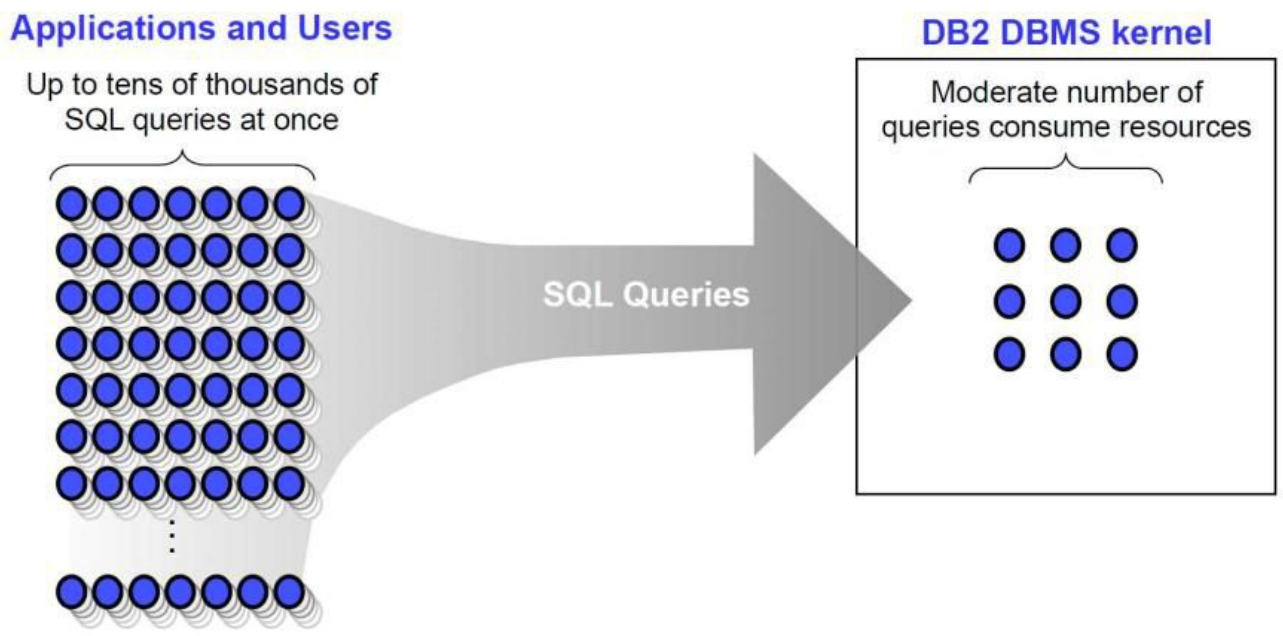
• ENTFÄLLT: Tune I/O

• ENTFÄLLT: Add Optimizer hints

• ENTFÄLLT: Statistics collection

Im Prinzip ist also bei BLU nur noch ein "Load and Go" notwendig: Der Admin muss nur sein Datenmodell kennen und wissen, welche Tabellen er umwandeln will.

Der Workload Manager



Der Workload Manager verteilt die Queries so auf die Cores, dass alle Kerne gut ausgelastet sind.
Foto: IBM

Die Anfragen an den Datenbank-Server steuert in DB2 BLU, wie in früheren Versionen, der Workload Manager. Das Tool ist bei analytischem Workload immer automatisch eingeschaltet und arbeitet unter anderem die Queries, die von den Applikationen an die Daten geschickt werden, in koordinierter Form ab.

Die Grafik veranschaulicht seine Aufgabe: Der Workload Manager verteilt die Queries so auf die Cores, dass alle Kerne gut ausgelastet sind und dadurch eine effiziente und koordinierte Abarbeitung der Queries erfolgt.

Gesteuert wird der Workload Manager durch die Umgebungsvariable

DB_WORKLOAD = ANALYTICS

Sobald diese Variable aktiviert ist, ist DB2 per Default für BLU optimiert.

Zu guter Letzt der Hinweis, dass die Beschleunigungstechnologie BLU die Performance nicht in allen Datenbank-Bereichen verbessert. Business-Transaktionen mit vielen schreibenden Zugriffen, wie sie typischerweise bei ERP-Systemen oder OLAP-Datenbanken vorkommen, sind kaum leistungsstärker als vorher - aber das gilt nicht nur für BLU, sondern für alle spaltenorientierten Datenbanken. Bei Lesezugriffen in großen Datenbanken im Data-Warehouse-Umfeld hingegen spielt BLU seine Fähigkeiten voll aus.

DB2 BLU soll ab dem zweiten Quartal 2013 zunächst für AIX und Linux verfügbar sein.

Links im Artikel:

[1] **http://wwwdb.inf.tu-dresden.de/birte2011/files/keynote.pdf**

IDG Tech Media GmbH
Alle Rechte vorbehalten. Jegliche Vervielfältigung oder Weiterverbreitung in jedem Medium in Teilen oder als Ganzes bedarf der schriftlichen Zustimmung der IDG Tech Media GmbH. dpa-Texte und Bilder sind urheberrechtlich geschützt und dürfen weder reproduziert noch wiederverwendet oder für gewerbliche Zwecke verwendet werden. Für den Fall, dass auf dieser Webseite unzutreffende Informationen veröffentlicht oder in Programmen oder Datenbanken Fehler enthalten sein sollten, kommt eine Haftung nur bei grober Fahrlässigkeit des Verlages oder seiner Mitarbeiter in Betracht. Die Redaktion übernimmt keine Haftung für unverlangt eingesandte Manuskripte, Fotos und Illustrationen. Für Inhalte externer Seiten, auf die von dieser Webseite aus gelinkt wird, übernimmt die IDG Tech Media GmbH keine Verantwortung.