

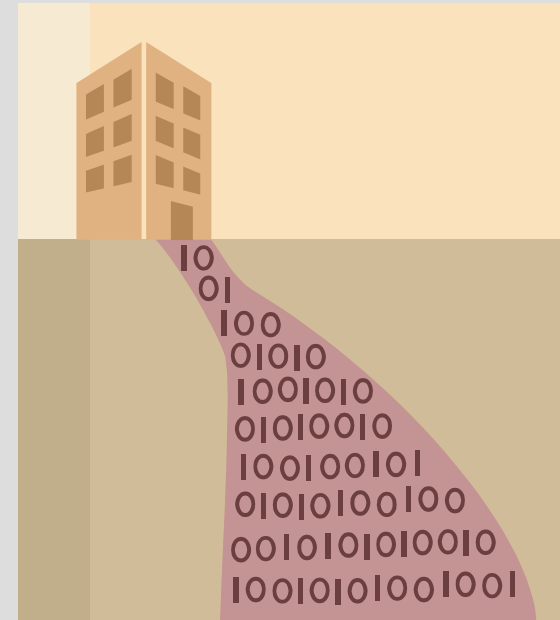
Das Zettabyte

CeBIT 2011

Dr. Wolfgang Martin
Analyst, iBonD Partner und Ventana Research Advisor

Das Zetabyte: analytische Datenbanken

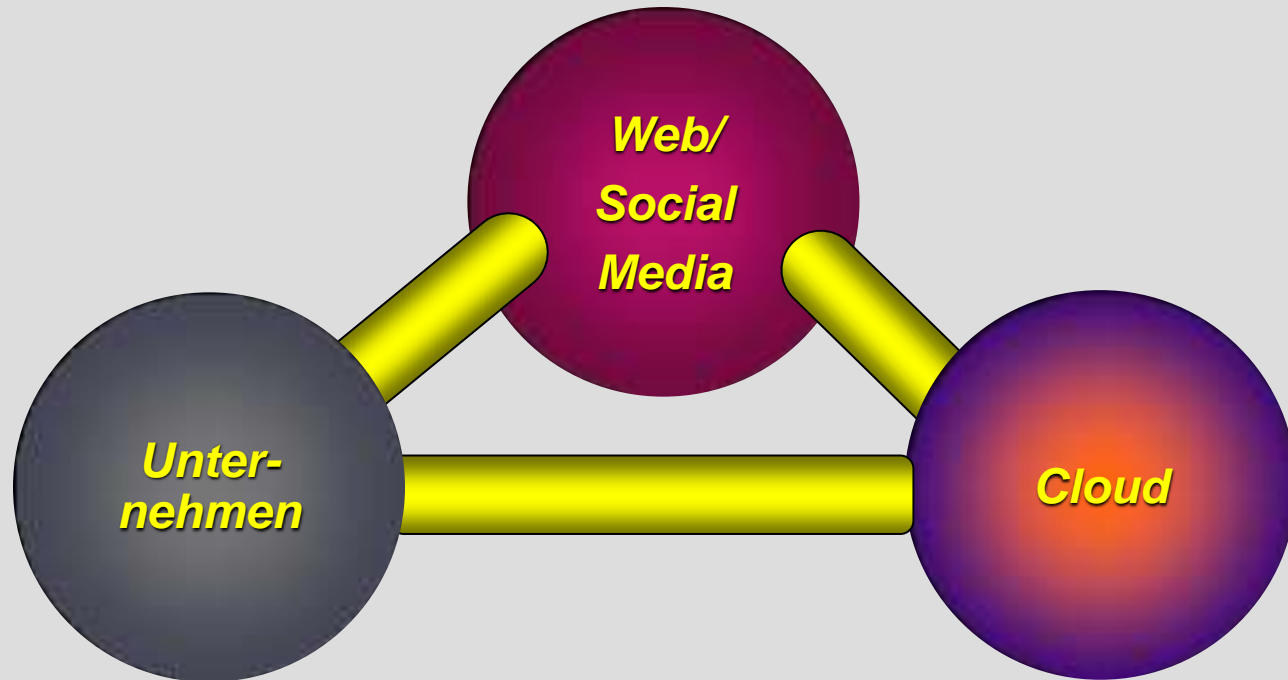
- Die Datenflut.
- Analytische Datenbanken: Was ist neu?
- Analytische Datenbanken: Die Marktspieler.



Die Zahl

- 1 B 1
- 1 KiloB 1,000
- 1 MegaB 1,000,000
- 1 GigaB 1,000,000,000
- 1 TeraB 1,000,000,000,000
- 1 PentaB 1,000,000,000,000,000
- 1 ExaB 1,000,000,000,000,000,000
- 1 ZettaB 1,000,000,000,000,000,000,000

Die Datenflut

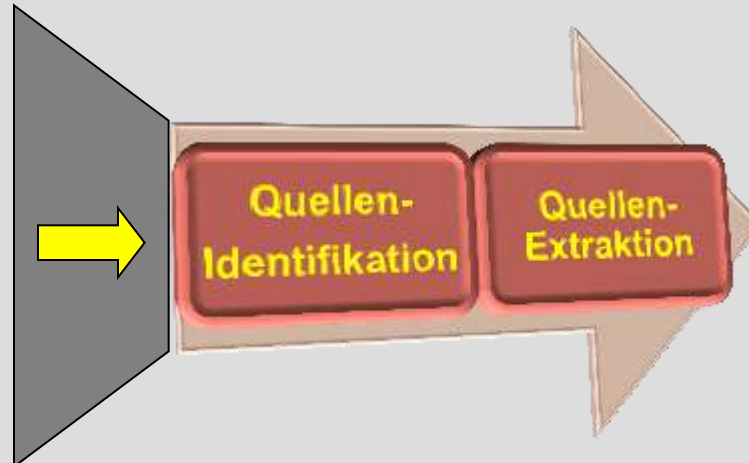


Mehr und mehr Daten, strukturiert und unstrukturiert.
Wie lässt sich das Potenzial erschließen?

Beispiel: Web als Datenquelle für CRM

Events als Auslöser

(neue Marktspieler, neue Produkte, Kampagnen, Neuigkeiten, etc)



***semantische
Web-Crawler***

Soziale Medien

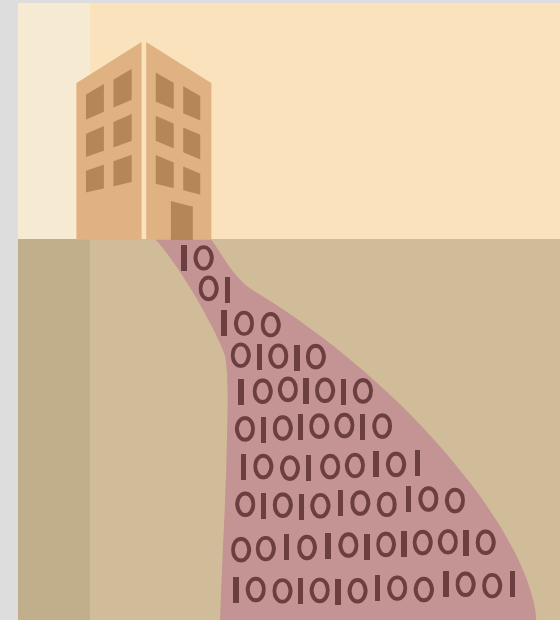
(Blogs, Wikis, Podcasts, Foren, Newsgroups, Second Life, Search, Tagging, etc)

Teilnehmer

(Unternehmen, Kunden, Konsumenten, Presse, Organisationen, Parteien, Institute etc)

Das Zetabyte: analytische Datenbanken

- Die Datenflut.
- Analytische Datenbanken: Was ist neu?
- Analytische Datenbanken: Die Marktspieler.



Analytische Datenbanken

Die Treiber:

- Das Datenvolumen steigt schneller als die Leistung von traditionellen Datenbanken.
- Man schafft es einfach nicht mehr, Daten im Detail zu analysieren, da es zu lange dauert.

Gartner sagt im „Magic Quadrat for Data Warehouse Database Management Systems 2010“: „Gartner-Klienten stehen bei der Abfrage von Data Warehouses immer häufiger vor erheblichen Performanceproblemen. Auf Grundlage dieser Informationen dürften rund 70 % aller Data Warehouses mit derartigen Problemen zu kämpfen haben.“

Konzepte analytischer Datenbanken

- Spalten-Orientierung
- Daten-Komprimierung
- Zugriffsmethoden und -Algorithmen
 - Google: „MapReduce“
- Parallelisierung
- In-Memory-Verarbeitung
- Data Appliances

Keine Wunder erwarten

- Eine analytische Datenbank macht ein physikalisches Datenbankdesign und Tuning weitgehend obsolet.
- Sie ersetzt aber keineswegs das logische, fachliche Design der analytischen Datenbank.
- In diesem Sinne bleibt weiterhin ein Information Management auch mit analytischen Datenbanken unabdinglich.
 - Stamm- und Metadaten-Management, ein Datenqualitäts-Management, eine Information Governance und die anderen Aufgaben im Information Management bleiben auch mit analytischen Datenbanken kritische Erfolgsfaktoren.

Noch eine Datenbank?

- Eine analytische Datenbank ersetzt nicht die herkömmlichen Datenbanken in der Transaktionsverarbeitung. Analytische Datenbanken sind eine neue Generation von Datenbanken für analytische Aufgaben im Unternehmen. Ein Unternehmen braucht heute eben in der Regel zwei unterschiedliche Datenbanktechnologien.
- Keine Regel ohne Ausnahme: Oracle Exadata und SAP Business By Design.

Analytische Datenbanken

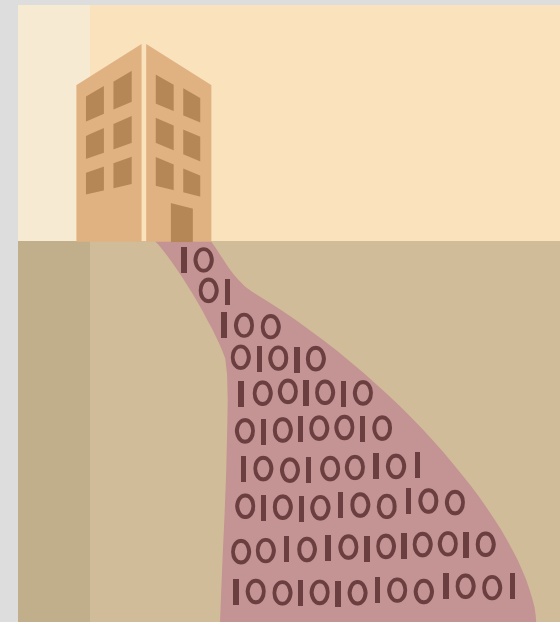
Vorteile analytischer Datenbanken:

- Informationen sind flexibler abrufbar und stehen bis zu 100mal schneller zur Verfügung.
- Die Nutzerzufriedenheit erhöht sich signifikant aufgrund des schnelleren und flexibleren Zugriffs auf Information. Es können jetzt Daten analysiert werden, die vorher ohne Nutzen, aber mit Kosten gespeichert wurden. Das unterstützt und schafft bessere Entscheidungen.
- Die IT wird entlastet, da die analytischen Datenbanken hoch automatisiert sind und ein spezielles Wissen über Datenbankdesign und Tuning deutlich weniger gefragt ist.



Das Zetabyte: analytische Datenbanken

- Die Datenflut.
- Analytische Datenbanken: Was ist neu?
- **Analytische Datenbanken: Die Marktspieler.**



Die Marktspieler

- Typ 1: parallelisierte herkömmliche Datenbanken, die in der Regel als Appliance angeboten werden, also eine spezielle Hardware und den parallelen Zugriffsmethoden und Algorithmen. Dabei sind solche Datenbanken dann immer noch zeilenorientiert.
- *Beispiele: EMC Greenplum, Ingres VectorWise, Kognitio, IBM DB2 (InfoSphere Warehouse), IBM Smart Analytics System, IBM Netezza, Oracle Exadata, SAS Scalable Performance Data Server, Teradata*

Die Marktspieler

- Typ 2: spaltenorientierte analytische Datenbanken, die weitgehend Hardware-unabhängig eingesetzt werden können.
- *Beispiele: Apache Hadoop HBase, HP Vertica, Illuminate, InfoBright, kx Systems, Sand Analytics, SAP Sybase IQ, Vectornova*
- *Bemerkung: Apache und InfoBright sind OpenSource-Lösungen.*

Die Marktspieler

- Typ 3: spaltenorientierte Datenbanken, die als Appliance teilweise mit spezieller Hardware angeboten werden, aber insbesondere In-Memory einsetzen.
- *Beispiele: 1010Data, Exasol, IBM Smart Analytics Optimizer, ParAccel, SAP HANA*
- Typ 4: Besondere Verfahren wie beispielsweise „Database Images“ (*Panoratio*) und OO-Datenbanken (*InterSystems*)



- Es gibt auch bereits erste Angebote von analytischen Datenbanken in der Cloud.
- *Beispiele:*
 - Die *Oracle Exadata Database Machine* kann in Private Clouds betrieben werden,
 - Anbieter wie *1010Data* und *Kognitio* bieten sowohl eine on premise-Lösung als auch ein DWaaS (Data Warehouse as a Service).

Analytische Datenbanken bringen den Nutzern ganz neue Möglichkeiten, sowohl in der Skalierbarkeit, der Performance als auch in den Betriebskosten.

Neueste Techniktrends sind hierbei die Spaltenorientierung inklusive Komprimierung und Zugriffsverfahren, wie massiv parallele Verarbeitung sowie die In-Memory-Technologie.

Wir meinen: Eine Evaluierung lohnt sich. Damit sollte man nicht mehr warten! Ein Business Case ist schnell gefunden.



White Paper zu Analytik und Performance Management (Kostenfreier Download)

http://www.wolfgang-martin-team.net/BI-BPM-SOA_dt.php

Kontakt:

wolfgang.martin@wolfgang-martin-team.net