

Big Data (Warehouse?)

Peter Welker
Trivadis GmbH
Stuttgart

Einleitung

Was ist dieses Big Data eigentlich? Wer braucht es? Wie unterscheidet es sich vom klassischen Data Warehousing? Handelt es sich womöglich nur um den schieren Zuwachs von Daten altbekannter Strukturen oder wird die Welt auch hier komplexer? Wie verarbeitet man das am besten?

Dieser Artikel grenzt traditionelle DWH-Szenarien von »neuen«, mit herkömmlichen Mitteln nur unzureichend lösbaren Aufgaben im Schlagschatten der modernen Datenflut ab und betrachtet dabei beide »Welten« und wie sie diesen Herausforderungen begegnen.

Was ist das eigentlich: „Big Data“?

Keine IT-Konferenz ohne »Big Data«. Auch diesmal nicht, sorry! Bei der letzten DOAG BI gab es einen eigenen Track zu diesem Thema und auf der TDWI Messe im Juni fanden sich 7 Vorträge dazu – nicht gerechnet die Präsentationen, die dieses Thema zumindest streiften. Und geht es nach den Analysten, ist der Hype gerade erst angelaufen.

Aber was ist eigentlich die Ursache für diese Welle? Das Datenvolumen in IT Systemen steigt doch schon immer kontinuierlich an.

Geschätzte **1,8 Zettabytes** (oder rund 1800 Exabytes) an Daten hat die Menschheit im Jahr 2011 weltweit erzeugt und repliziert – dazu gehört auch deren Verbreitung im Internet. Das ist neun Mal so viel wie noch 5 Jahre zuvor [1]. Stellte man die für die Speicherung dieser Daten benötigten Server-Festplatten hochkant nebeneinander, reichte dieser Gürtel zweimal um den Äquator. Wirklich neu produziert und mehr oder weniger dauerhaft gespeichert wurden 2010 immerhin noch **13 Exabytes** Daten. Das McKinsey Global Institute (MGI) sieht in der Nutzung der in diesen Daten enthaltenen Informationen alleine im amerikanischen Gesundheitswesen ein jährliches Potential von rund **300 Milliarden USD**. Weitere **250 Milliarden €** sind es gemäß dieser Studie in Europas öffentlichem Sektor [2].

Das sind beeindruckende Zahlen. Verpassen wir also etwas, wenn wir nicht umgehend in neuen Speicherplatz investieren? Dieser Quantensprung an Quantität soll nicht darüber hinwegtäuschen, dass auch heute nur ein winziger Teil dieser Informationen für die Entscheidungsfindung wirklich relevant sind. Werfen wir einen Blick auf die Zusammensetzung dieser Datenflut.

Un-/Poly-/Semi-/Irgendwie strukturierte Daten

Ein zunehmend großer Anteil der „neuen“ Daten basiert auf weitgehend unstrukturierter, aber immens umfangreicher Formate wie Bilder, Audio, Video oder Layout. Auf YouTube wurde im Frühjahr 2010 jede Minute 24 Stunden an Videomaterial hochgeladen, heute sind es über 50. Daneben steigt auch der Platzbedarf innerhalb desselben Mediums drastisch. Ein Full-HD-Video mit zeitgemäßer Tonspur benötigt mit bis zu 6,75 Megabytes pro Sekunde leicht 30 Mal so viel Speicherplatz wie derselbe Film im guten, alten PAL Fernsehen, transportiert aber für den überwiegenden Teil aller denkbaren Anwendungsfälle immer noch dieselben Informationen. Und mit 4K2K und 8K4K stehen bereits die

Kinder und Enkel mit vierfacher beziehungsweise sechszehnfacher Auflösung in den Startlöchern. Abgesehen von wenigen strukturierten Anteilen (Tags) sind solcherart Daten für die Unternehmenssteuerung aber doch eher irrelevant. Das mag sich gegebenenfalls durch Anwendungen wie „Automatisierte Gesichtserkennung“ o.Ä. in der Zukunft partiell ändern, wird sich aber selbst dann wohl eher in der Aggregation dieser Daten in kleine, klar strukturierte Einheiten niederschlagen. Für den Rest genügt »dummer« Speicherplatz ohne komplexe Suchmöglichkeiten irgendwo in der Cloud.

Interessanter ist das Potential für uneinheitlich-strukturierte Daten wie den personifizierbaren, durch Volltext und Multimedia angereicherten „Social-Media Content“ aus facebook, twitter & Co. Dieser ist zwar mit klassischen BI Ansätzen (noch) nicht leicht zu analysieren, bietet aber bereits heute die Möglichkeit, Meinungsführer und Trends zu identifizieren und wird zunehmend fürs Marketing genutzt.

Besonderes Augenmerk benötigen auch die Informationen, die sich immer mehr durch permanent von Maschinen produzierte Daten ergeben. Man spricht auch medienwirksam vom „Internet of Things“. Dazu darf man sowohl geografisch/persönliche Ereignisse wie die Lokalisierung von Smartphones und das Lesen von RFIDs, als auch die heute fast schon klassischen Web-Logs oder die immer mitteilungsüchtigere Produktionssensorik zählen.

Strukturierte Daten

Natürlich spielen auch „herkömmliche“ Daten eine bedeutende Rolle für das Datenwachstum. Unternehmen und Behörden erzeugen heute deutlich größere Informationseinheiten, speichern also immer mehr Informationen zu jedem Kunden und mehr Details zu jeder Transaktion. Gleichzeitig wächst auch die Anzahl der Informationseinheiten und damit die Granularität in Stammdaten und Bewegungsdaten gleichermaßen.

Diese Daten bilden nach wie vor den Kern des traditionellen BI – und sind nicht selten über bestimmte Zeiträume hinweg zur Gewährleistung höchster Transparenz erforderlich, wie zum Beispiel Diagnose- und Leistungsdaten im Gesundheitswesen, Call Data Records bei den Telekommunikationsdienstleistern oder Kontenbewegungen bei Banken.

Einheitliches Kundenverständnis

Durch das Zusammenführen aller Informationen zu einem einheitlichen Kundenverständnis, entstehen wertvolle zusätzliche Erkenntnisse.

Nun stellt sich die Frage, wie mit dem **Dilemma des Datensammelns** umgegangen werden soll:

Muss ich alle Daten über lange Zeit sammeln, um schon heute für zukünftige Anforderungen gerüstet zu sein? Oder genügt es, diese neuen Datenquellen erst anzuzapfen, sobald konkrete Anforderungen aufkommen?

Letzteres erlaubt deutlich geringeren Investitionen, allerdings auf Kosten des Risikos, von neuen Erkenntnissen nicht schnell genug profitieren zu können und auf deren „Geschichte“ zu verzichten.

Big Data, eine „disruptive“ Technologie?

Die große Riege der NoSQL-, BigTableClone- und sonstigen verteilten „Datenbanken“ – also Hadoop, Dynamo, Cassandra und Co. haben ihre Einsatzfähigkeit für (uneinheitlich) strukturierte Daten in Einzelfällen längst eindrucksvoll unter Beweis gestellt. So betreibt bspw. Yahoo! Hadoop Cluster mit mehreren hundert Petabytes an Daten.

Wenn Amazon, Yahoo, Facebook und Co. mit diesen Technologien die größten Datenmengen erfolgreich verarbeiten, stellt sich natürlich auch für jeden von uns die Frage nach der richtigen Technologie. Bin ich mit meiner aktuellen BI-Lösung auf dem Holzweg? Wird „Big Data“ die Art und Weise, wie wir mit unseren entscheidungsrelevanten Daten umgehen von Grund auf verändern?

Die Antwort ist zunächst schlicht »**Nein**«. Heute deutet alles darauf hin, dass die gängigen Big Data Ansätze mit Hadoop & Co. nicht geeignet sind, Aufgaben der klassischen BI 1:1 zu übernehmen.

Warum? Die Nutzung von Hadoop mittels seines Map Reduce Frameworks ist zwar prinzipiell extrem flexibel – schließlich kann ich alles tun, was mittels Programmierung und Dateien umsetzbar ist – dadurch aber auch mit sehr hohem Entwicklungsaufwand verbunden und direkt für Endanwender praktisch nicht nutzbar. Das entspricht in etwa dem Unterschied von modernen Datenbanken und modernen Filesystemen. Das eine ist kein Ersatz für das andere, denn die Aufgabenstellungen sind schlicht unterschiedlich. Mit einfachen Files ist mein Spielraum unbegrenzt. Ich habe alle Freiheiten. Ich habe allerdings auch ungeheure Aufwände, um in jedem meiner Programme die Funktionalität einer typischen Datenbank nachzubilden.

Eine interessante Zwischenlösung stellt eine Datenbank dar, die Hadoop als Filesystem nutzt. Das tut beispielsweise Hive, eine Art SQL-Schnittstelle für Hadoop, das Map Reduce Zugriffe aus SQL-Befehlen generiert. Solche Lösungen sind aber funktional noch sehr weit von der Klasse traditioneller Datenbanksysteme entfernt - und letztlich steigern auch sie die Einfachheit wieder auf Kosten der Flexibilität.

Letztlich ist der Anspruch der BI an Einfachheit für Entwickler und Anwender und die Ausrichtung auf seit bald 20 Jahren etablierte Methoden wie Pivottisierung und dimensionale Modelle im Zusammenspiel mit extrem geringen (Sub-Second) Latenzen von MOLAP Technologien und ausgefeilten Aggregationsfeatures zu ausgeprägt, um einfach durch eine Technologie ersetzt zu werden, deren Stoßrichtung schlicht Performance bei sehr großen Datenmengen, Kosten und Flexibilität ist. Langfristig dürften die BI Anwenderwerkzeuge Datenquellen wie Hadoop & Co. als Alternative für bestimmte Aufgaben einbeziehen. Solange lassen sich mit Big Data Ansätzen aber eher gänzlich neue Aufgabenfelder erschließen oder zweckentfremdende Prozesse aus bestehenden Data Warehouse eliminieren. Besonders aber kann die Vor- und Aufbereitung entscheidungsrelevanter Informationen aus bisher unzugänglichen Daten den etablierten BI-Systemen schon heute gute Dienste leisten. Die neuen Technologien sind also nicht „disruptiv“. Sie ergänzen sich vielmehr glänzend.

Lösungsansatz „Smart Data“

Die hohen Anforderungen an die Analyse strukturierter Daten können heute mit entsprechenden DWH Appliances bis in den zwei- und dreistelligen Terabyte-Bereich hinein bewältigt werden – allerdings zu entsprechenden Kosten. Die große Masse der uneinheitlich strukturierten Daten werden hingegen mit entsprechenden Lösungen aus dem NoSQL & BigTable Umfeld (bspw. Hadoop) verwaltet, selektiert und in bestimmtem Umfang auch analysiert. Bei Bedarf können die dort gehaltenen Daten dann aufbereitet und ins klassische DWH geladen werden. Aus BigData würde somit für BI-relevante Anwendungsfälle „SmartData“.

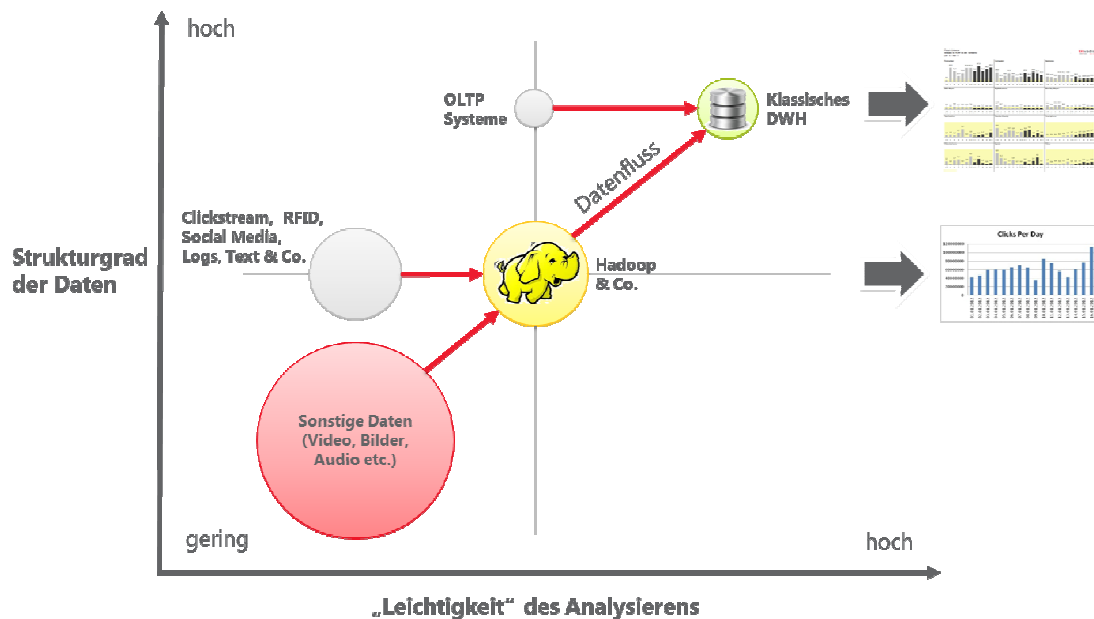


Abbildung 1: Mehrstufiges Konzept zur Analyse von Daten unterschiedlichen Strukturierungsgrades

BigData und DWH Appliances

Inzwischen stellen einige Vendors bereits zwei verschiedene Varianten von Appliances zur Auswahl, BigData Lösungen und MPP¹-Data Warehouse Cluster Lösungen. Die für „BigData“ deklarierten Lösungen wie Oracle BigData Appliance oder EMC Greenplum HD bauen weitgehend auf Hadoop auf und stellen dafür spezielle Hadoop Distributionen mit zahlreichen Erweiterungen und eigenen Implementierungen beispielsweise von Key-Value Stores (Teradata Aster) zur Verfügung.

Wir wollen uns im Zweiten Teil des Artikels allerdings auf die Data Warehouse Appliances konzentrieren und betrachten, welche Lösungen hier angeboten werden und welche Architekturen dafür zum Einsatz kommen. Typische Vertreter dieses Genres sind Oracle Exadata, Teradata, IBM Netezza, HP Vertica, EMC Greenplum, Microsoft PDW oder Exasol.

Diese Data Warehouse Appliances fokussieren sich auf klassische DWH Methoden und nutzen dafür spezielle RDBMS Implementierungen mit weitgehend standardisierter SQL Schnittstelle. Das Besondere daran: Es handelt sich um Datenbank-Cluster, die nach dem Scale-Out Prinzip große Datenmengen auf einzelne, relativ kleine Serverknoten verteilen. Durch geschickte „Partitionierung“ der Aufgaben werden die Daten dann möglichst direkt auf den einzelnen Knoten vorverarbeitet, also beispielsweise gefiltert, gejoined und aggregiert. Abschließend werden die stark verkleinerten Resultate an einer Stelle zusammengeführt und dem Anwender zurückgegeben. Dafür müssen nur relativ kleine Ergebnismengen über das interne Clusternetzwerk geschickt werden.

¹ Massive Parallel Processing

Generelle Aufgabenstellungen bei großen Datenmengen

Prinzipiell sind die Aufgabenstellungen an Skalierung durch Verteilung bei allen Clusterlösungen, ganz gleich ob HDFS oder Datenbank-Cluster, sehr ähnlich. Es geht dabei vereinfacht gesagt um Folgendes:

- Eine einzelne Aufgabe auf vorhandene Ressourcen zu verteilen (Partitionierung)
- Diese Verteilung möglichst gleichmäßig durchzuführen (Skewing vermeiden)
- Auch bei mehrstufigen Aufgaben, verfügbare Ressourcen immer optimal auszulasten (Pipelining)
- Dabei möglichst viele Operationen innerhalb eines Knotens auszuführen (Lokalität)

Dabei gibt es natürlich jede Menge Möglichkeiten, die Auslastung der Ressourcen gering zu halten bzw. optimal aufeinander abzustimmen, um gute Antwortzeiten zu erreichen wie bspw:

- Organisation (Spalte/Zeile) und Kompression der Daten, um einen höheren DeFacto Durchsatz zu erzielen
- Maßnahmen zur Reduktion der zu verarbeitenden Daten wie Indexierung, horizontale und vertikale Partitionierung, Vorhalten von Min/Max-Werten für Datenbereiche usw.
- Priorisierung bestimmter Abfragen nach Anwendergruppe oder zu erwartender Laufzeit (Ressourcenmanagement) etc.

Da analytische Aufgaben meist eine ganze Reihe von elementaren Operationen beinhalten (Scans, Joins, Filter, Gruppierungen, Sortierungen), kommt es auch und besonders darauf an, wie der Query-Optimizer eine Aufgabe zerlegt, wie er also die einzelnen Schritte zur Erfüllung der Aufgabe auswählt und kombiniert. Soll er beim Join die Daten einer Tabelle an alle Knoten schicken oder nur Teile davon bestimmte Knoten? Soll er zuerst die Tabelle A oder die Tabelle B scannen? Um diese und noch viel mehr Fragen beantworten zu können, muss bekannt sein:

- Nach welchen Kriterien werden die Daten auf den Knoten verteilt (Distribution)
- Welche Tabelle hat welche Datenmengen und welche Daten befinden sich in den einzelnen Spalten (Statistiken)

Beispiele für Lösungsansätze

Generell sind die Anbieter mit Ihren Lösungen auf die Anforderungen eingestellt. Alle nutzen umfangreiche **Statistiken** über die Datenverteilung für die Ermittlung von **Ausführungsplänen** und bieten Lösungen zur Reduktion des Datenverkehrs über den Interconnect.

Zur **Verteilung/Distribution** der Daten auf den einzelnen Knoten teilen sich die Lösungen in zwei Lager: Die einen nutzen eine klassische SharedNothing-Architektur, haben also pro Knoten lokale Platten und verteilen die Daten einer Tabelle zufällig oder nach dem Hashwert von definierbaren Spalten möglichst gleichmäßig auf die Knoten. Die anderen – eigentlich nur Oracle Exadata – fahren einen gemischten Ansatz. Dabei werden zwei Typen von Knoten genutzt: Die einen (klassische DB Knoten, meist als RAC konfiguriert) sehen die anderen (Storage Server) als eine Art SharedDisks. Auf den Storage Servern wiederum werden die Daten dann aber gleichmäßig verteilt. Dieser eher hybride Ansatz erlaubt Oracle die 1:1 Nutzung seiner Standard-Funktionalität. Natürlich speichern alle

Anbieter standardmäßig alle Daten redundant auf mehreren Knoten, um Ausfallsicherheit zu gewährleisten.

Alle erlauben auch die **Kompression** von Daten. Die Methoden dafür sind allerdings sehr unterschiedlich. Oracle Exadata, EMC Greenplum und mittlerweile auch Teradata bieten beispielsweise verschiedene Varianten (zeilenorientiert, spaltenorientiert) in verschiedenen Stärken an. Die rein spaltenorientierten Lösungen wie HP Vertica und Exasol komprimieren natürlich rein spaltenorientiert, was je nach Datenlage sehr effizient sein kann und tun dies „out of the box“. IBM Netezza hingegen nutzt spezielle Hardware (FPGAs), um beispielsweise die Kompression auf Disk möglichst effektiv zu gestalten.

Beim **Ressourcenmanagement** können nur wenige Anbieter direkt Einfluss auf die einzelnen Elemente einer Operation nehmen, also bspw. CPU, Disk und Netzwerk prozentual auf einzelne Aufgaben verteilen, um so bspw. „taktischen“ Anwendern eine garantierte Verfügbarkeit zu gewährleisten. Teradata hat hier sehr umfangreiche Möglichkeiten und Oracle Exadata bietet – im Gegensatz zur „normalen“ DB – neben der prozentualen CPU-Zuteilung auch eine prozentuale IO-Zuteilung an. Bei den meisten anderen Lösungen beschränken sich die Möglichkeiten auf die Nutzung von unterschiedlichen Queues pro Anwendergruppe oder vergleichbare Ansätze.

Standby, Backup & Co.

Was bedeutet die Einführung einer solchen Lösung nun aber aus systemadministrativer Sicht?

Backups kann man natürlich bei allen Systemen ganz klassisch (auch im Betrieb) ziehen. Aber ein „Snapshot“ wie man ihn aus der SAN Welt kennt ist bei lokalen Disks nicht möglich. Will man eine DB also klonen heißt es schlicht „kopieren“.

Bei den **Hochverfügbarkeits**-Lösungen (HA) wird es uneinheitlich. Manche erlauben, den Cluster auseinander zu ziehen (Stretch-Cluster). Das ist zwar bei relativ kleinen Abständen (bspw. 100m über eine Feuerschutzwand) noch performant möglich, bedingt aber auch eine Verdopplung der Redundanz, wenn nach dem Ausfall einer Seite die andere alleine nicht ohne Redundanz der Disks weiterlaufen soll. Wenn die Anforderungen allerdings eine Verteilung auf weit entfernte Standorte nötig machen (bspw. 100km), kann ein Stretch-Cluster auch sehr ineffizient werden, denn einerseits steigt mit der Entfernung naturgemäß die Latenz und andererseits sind optimierte Interconnects in WANs nicht mehr so ohne weiteres realisierbar.

Oracle bietet immer die Möglichkeit, eine **Standby** DB aufzubauen (DataGuard). Teilweise ist bei anderen Anbietern eine ähnliche Lösung vorhanden oder der Einsatz von Replikationssoftware möglich. Schlimmstenfalls muss man alle ETL Prozesse aber an zwei Standorten laufen lassen.

„Kraft“ versus „Intelligenz“

Neben den ganzen brachialen Möglichkeiten der Skalierung und der optimalen Ressourcennutzung darf man aber auch die „Intelligenz“ einer Lösung nicht vernachlässigen. So ist es beispielsweise schlicht Geldverschwendung, wenn man in einer typischen DWH Lösung seine Daten ausschließlich auf dem untersten Granularitätslevel analysiert, obwohl eine redundante Haltung voraggrierter Daten 80% der Abfragen mit einem Tausendstel des Aufwandes bedienen könnte.

Literaturverzeichnis

[1] IDC Reportserie: "Digital Universe", 2007 ff. Sponsored by EMC

[2] MGI Report: "Big Data. The next frontier for innovation, competition, and productivity." 2011

[3]

http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/de//archive/mapreduce-osdi04.pdf

Kontaktadresse:

Peter Welker
Trivadis GmbH
Industriestraße 4
D-70565 Stuttgart

Telefon: +49 (0)711-903 63 230

Fax: +49 (0)711-903 63 259

E-Mail peter.welker@trivadis.com

Internet: www.trivadis.com