

BUSINESS INTELLIGENCE UND BIG DATA GEHÖREN ZUSAMMEN

Was soll aus den bisherigen BI-Anwendungen werden, wo die IT-Welt doch derzeit nur noch über Big Data spricht? Die Praxis zeigt, dass beide Welten relevant sind und miteinander existieren können. Business-Intelligence-Umgebungen können durch die Analyse von Big Data profitieren. Dennoch werden BI-Organisationen vieles lernen und überdenken müssen.

von Sascha Alexander

Das Interesse an Big Data ist enorm und wird durch die Digitalisierung von Prozessen und des Datenaustauschs zwischen Menschen, IT-Systemen und Geräten sowie zwischen Geräten untereinander weiter gesteigert (Stichwort „Internet of Things“ IoT). Kaum ein Tag, an dem nicht über neue Möglichkeiten und Herausforderungen im Zusammenhang mit der Erfassung, Speicherung und Analyse großer, polystrukturierter Datenmengen zu lesen ist. Weniger thematisiert wird dabei, wie sich die Nutzung von Big Data zur bisherigen Business Intelligence (BI) verhält. Doch gerade BI-Organisationen müssen angesichts der hohen Investitionen der vergangenen Jahre und des dabei aufgebauten BI-Know-hows wissen, was dieser Trend für ihre künftige Arbeit bedeutet. Sind beispielsweise massive organisatorische und technische Veränderungen nötig, um Big Data und neue Analysetechniken einzubeziehen? Läuft es auf eine Koexistenz von Big-Data-Umgebungen und klassischen BI-Systeme hinaus, können sich beide Welten befruchten oder ist das Ergebnis gar eine Verdrängung vorhandener Systeme?

Aus der Beratungs- und Projektpraxis können wir sagen, dass Business Intelligence mit seinen typischen Disziplinen Reporting und Analyse heute wichtiger denn je ist, um Unternehmen zu steuern und das Geschäft vorausschauend zu planen. Die typische, allen vertraute BI-Architektur, in deren Mittelpunkt das Data Warehouse (mit seinen Schichten „Staging Area“, „Cleansing“, „Core“, „Data Mart“) steht und auf das dann die BI-Frontends zugreifen, wird uns noch lange begleiten. Doch ebenso deutlich zeigt sich, dass die Nutzung von Big Data und dem damit einhergehenden Einsatz von Analytics neue, von der klassischen BI nicht adres-

sierte „Use Cases“ ermöglicht. Es bedarf einer anderen Datenarchitektur, welche die Bewirtschaftung, Analyse und Aufbereitung poly- und unstrukturierter Daten (Sensordaten, Weblogs, Texte etc.) besser oder überhaupt erst unterstützt.

Typische Anforderungen an Big-Data-Analytics-Umgebungen sind die Datenaktualisierung in Echtzeit/Near Real-time/Batch, verbunden mit der hochparallelen Datenverarbeitung auch großer Datenmengen gegebenenfalls per „Streaming“ sowie die für Analytics typischen „fortgeschrittenen“ Analysen (statistische Verfahren, Methoden des Data Mining, Textmining). Bei der Referenzarchitektur in Abbildung 1 handelt es sich um eine sogenannte Lambda-Architektur, die vom ehemaligen Twitter-Softwareentwickler Nathan Marz entwickelt wurde. Ein wesentliches Merkmal ist das Konzept von unveränderlichen Daten (immutable data), d.h. dass Veränderungen an Daten nur Kopien dieser Daten erzeugen und die ursprünglichen Daten niemals verändert werden. In der Lambda-Architektur werden daher vorhandene Daten niemals aktualisiert. Jede Veränderung wird als eigenständiges Faktum angesehen, das zu einem bestimmten Zeitpunkt als wahr gilt. Eine Veränderung der Daten bedeutet also, dass nur ein neues Faktum für einen aktuelleren Zeitpunkt hinzukommt (append-only). Das zweite grundlegende Prinzip der Lambda-Architektur ist die Definition von Information als eine Funktion der Fakten. Informationen leiten sich also aus den Berechnungen der einzelnen Fakten her ab. Ziel ist es, eine Architektur aus skalierbaren Komponenten bei zugleich hoher Fehlertoleranz zu schaffen. →





Big Data Referenzarchitektur Lambda-Architektur

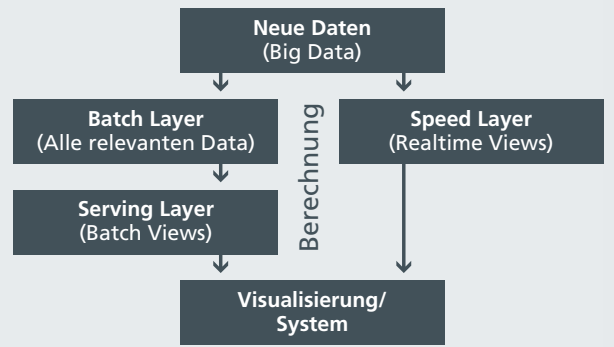


Abbildung 1: Die Aufbereitung und Analyse von Big Data ist mit einem anderen Architekturkonzept verbunden als in der klassischen BI. Quelle QUNIS

Unterscheidungsmerkmale BI und Big Data

	BI / Data Warehouse	Big Data
<ul style="list-style-type: none"> • Datenaktualisierung • Verarbeitungsgeschwindigkeit • Analytik • Datenmengen • Anwendungsszenarien 	<ul style="list-style-type: none"> • Täglich / mehrfach pro Tag • Primär bei der Beladung oder Abfrage • Einfach bis mittelkomplexe Berechnungen und Abfragen • Kleine bis mittelgroße Szenarien (mehrere TB) • Reporting/Analyse für Finanzen, Vertrieb, Personal, Marketing, IT, Einkauf etc. • Unternehmensplanung 	<ul style="list-style-type: none"> • Real Time / Near Time / Batch • Hochparallele Verarbeitung bis hin zum Streaming • Bis hin zu hochkomplexen analytischen Applikationen • Bis zu riesigen Datenmengen • Erweiterte Analytic, Streaming-analysen, Textmining etc.

Abbildung 2: Die Architekturen von BI- und Big-Data-Systeme wurden für unterschiedliche Anforderungen im Datenmanagement und -analyse geschaffen. Quelle: QUNIS

In der Big-Data-Lambda-Architektur hält ein „Batch Layer“ (beispielsweise HADOOP) sämtliche Fakten in redundanter Ausführung vor, übernimmt die Berechnungen und weist eine teilweise sehr hohe Latenz auf. Die Berechnungen werden deshalb im „Serving Layer“ gespeichert, was solche Latenzen verringert. Hier erfolgt zudem die Abfrage der gewünschten Informationen (externe Systeme). Im „Speed Layer“, der komplexesten Komponente, finden sich ebenfalls alle relevanten neuen Daten, die dort aber nur temporär für Berechnungen liegen und gelöscht werden, sobald neue Daten im Batch Layer verfügbar sind. Ein externes System wird dann die Ergebnisse aus Serving Layer und Speed Layer kombinieren, um eine aktuelle Sicht auf die Daten zu gewähren.

Business-Intelligence-Architekturen sind hierfür nicht geeignet und auch nie konzipiert worden. Ihre Domäne sind Szenarien, für die einmal oder mehrfach am Tag erfolgende Datenaktualisierungen genügen und in denen es vor allem um das Beladen vergleichsweise überschaubarer, historisierter Datenmengen sowie die Abfrage auf neuen (transaktionalen, strukturierten/modellierten) Daten geht (siehe Abbildung 2). Typische Anwendungen fallen in den Bereich des bisherigen Reportings und der Analyse (OLAP) in Fachbereichen wie Finanzen, Vertrieb, Personal, Marketing, IT oder Einkauf sowie der Unternehmensplanung.

Angesichts der unterschiedlichen Anforderungen werden daher Business-Intelligence-Systeme auch künftig in vielen bisherigen Anwendungsgebieten ihre Rolle behalten. Dies nicht zuletzt auch deshalb, weil hohe Investitionen in die bisherige BI-Infrastruktur getätigt wurden und dabei Standardprozesse etabliert und viel Know-how aufgebaut wurden. Zudem zeigt sich schon heute im Markt, dass traditionelle BI-Software-Anbieter ihre Plattformen um zusätzliche Schnittstellen und funktionale Erweiterungen anreichern, um so im Segment der Advanced Analytics mithalten zu können.

Business-Intelligence-Umgebungen können zugleich durch Big Data und die Methoden und Technologien für die hochskalierbare Erfassung, Speicherung und Analyse polystrukturierter Daten erheblich profitieren. Gerade im Finanzbereich besteht die Notwendigkeit, zusätzliche Geschäfts- und Marktinformationen in die Planung und Risikoanalyse einzubeziehen. Marketing und Vertrieb können ihre Kunden- und Auftragsdaten beispielsweise durch Churn-Analysen, eine Kundensegmentierung oder Analysen von Web-Logs um wichtige Informationen anreichern, um Portfolio und Preise zu optimieren.

Es kommt also künftig auf die richtige Kombination an, um von beiden Welten maximal zu profitieren. Richtig eingebunden kann die Nutzung von Big Data dafür sorgen, dass Prozesse optimiert, Märkte transparenter und neue Geschäftsmodelle möglich werden, indem ein Unternehmen seine Produkte und Dienstleistungen individualisieren kann. Eine sorgfältige technische und fachliche Planung sowie Anforderungsanalyse (Use Cases) sind dabei entscheidend fürs Gelingen. ●

Sascha Alexander

ist Manager Marketing & Kommunikation bei der QUNIS GmbH in Neubeuern. QUNIS ist auf die Beratung und Umsetzung von Projekten für Business Intelligence, Big Data und Analytics spezialisiert. Zuvor arbeitete Herr Alexander in gleicher Funktion beim Business Application Research Center (BARC) und war Chefredakteur des Portals und Magazins für Finanzvorstände „CFOworld“ sowie langjähriger Fachredakteur der „Computerwoche“. Sascha.alexander@qunis.de www.qunis.de

