

Einsatzszenarien für SAP HANA

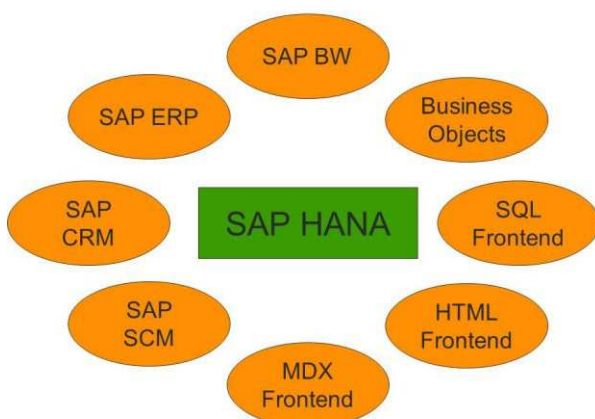
von Markus Alterauge



Management Summary

Unter dem Namen SAP HANA (High Performance Analytic Appliance) vermarktet die SAP eine Kombination aus Hardware und Software (Appliance). Herausragendes Merkmal der Appliance ist die enorme Geschwindigkeit der Datenbearbeitung und -bereitstellung, die SAP HANA als Beschleuniger bestehender Anwendungen qualifiziert, die aber auch vollkommen neue Möglichkeiten, wie beispielsweise im Bereich Echtzeitanalyse, erschließt.

Das vorliegende Whitepaper zeigt Einsatzszenarien und Potential von SAP HANA auf. Auch wenn die Mehrzahl der aktuellen Implementierungen im klassischen SAP Umfeld und hier insbesondere im Bereich Business Intelligence zu finden sind, so ist es wichtig festzustellen, dass der Einsatz von SAP HANA nicht auf ein SAP-Umfeld begrenzt ist. SAP HANA verfügt über eigene Modellierungswerkzeuge und eine eigene Scriptsprache. Diverse Schnittstellen ermöglichen den Datenaustausch und erlauben auch Nicht-SAP-Frontends, in gleicher Weise von den Geschwindigkeitsvorteilen zu profitieren. Eine integrierte Planungsebene sowie ein eigenes Sicherheitskonzept unterstreichen den agnostischen Charakter von SAP HANA.



SAP HANA – Basis für die SAP Lösungen der Zukunft und offen für hochperformante Lösungen im Nicht-SAP Umfeld

Ob in Bezug auf die Integration der ERP Suite, der Schaffung eines Cloudangebotes oder dem Aufbau einer mobilen Infrastruktur, im Umfeld der

SAP ist SAP HANA die zentrale Komponente, an der sich auf Sicht vieler Jahre alle neuen Entwicklungen orientieren werden¹.

„In-Memory“ - Begriffserklärung

Schalten Sie ihren Rechner ein, so sehen Sie sich im Allgemeinen mit der Rechnerarchitektur der zurückliegenden Jahrzehnte konfrontiert: Auf einem Massenspeicher, z.B. einer Festplatte, lagert Ihr gesamter Datenbestand. Teile dieses Datenbestandes werden nun in den Arbeitsspeicher geladen (Programme und Anwendungsdaten), die zentrale Recheneinheit (CPU) koordiniert diese Vorgänge. Bis der Rechner arbeitsbereit ist, vergeht eine gewisse Zeit.

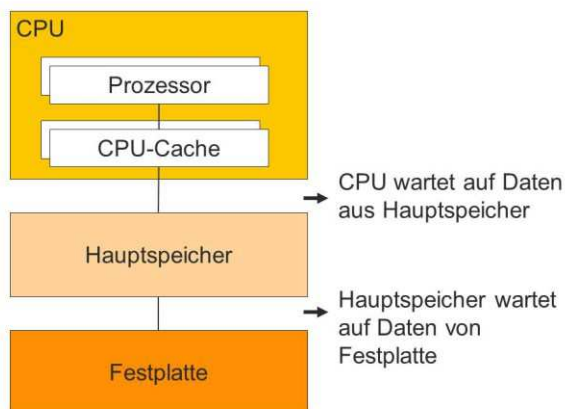
Im weiteren Verlauf Ihrer Arbeit greifen sie auf andere Anwendungen und Programme zu und sehen sich dabei immer wieder mit Wartezeiten konfrontiert. Diese Wartezeiten entstehen, wenn der CPU Daten fehlen. Dann fragt diese zunächst beim Arbeitsspeicher an, ob die benötigten Daten dort vorhanden sind. Werden sie dort nicht gefunden, geht eine Anfrage an den Massenspeicher und die Daten werden von dort in den Hauptspeicher und die CPU übertragen.

Das aufgezeigte Szenario aus dem Endanwenderumfeld ist in seiner Implikation auf das professionelle Umfeld übertragbar. Auch in großen IT-Umgebungen kommt es immer wieder zu Wartesituationen. Bestes Indiz hierfür ist die Auslagerung von rechenintensiven Operationen in lastarme Zeiten.

Die Geschwindigkeit der Datenverarbeitung wird wesentlich durch die Zugriffszeit auf Speicher bestimmt. Während diese bei handelsüblichen Festplatten um ca. 10 ms liegt, kann Arbeitsspeicher (RAM-Speicher) eine Anfrage in ca. 250 µs bedienen. Ein Weg, um die Geschwindigkeit der Datenverarbeitung zu erhöhen, ist demnach der Ersatz von Festplattenspeicher durch RAM-Speicher. In SAP HANA wurde dieser Ansatz so konsequent umgesetzt, dass die gesamte Datenbank nur

¹ Die Debatte um alternative Lösungen im In-Memory-Umfeld soll in diesem Whitepaper bewusst nicht aufgegriffen werden. Der Fokus liegt allein auf den Möglichkeiten von SAP HANA.

noch aus RAM-Speicher aufgebaut ist. Der gesamte Datenbestand steht also ständig „In-Memory“ zur Verfügung. Das klassische Terzett Festplatte-Arbeitsspeicher-CPU wird durch ein Duett von Arbeitsspeicher und CPU ersetzt. Die spaltenweise Datenablage und eine neue Architektur der Hardware optimieren die Zusammenarbeit zwischen CPU Cache und Arbeitsspeicher und minimieren dadurch Wartezeiten der CPU.



Typische Flaschenhalse klassischer Hardwarearchitektur

„In-Memory“ – von der Vision zur Realität

Schon heute entscheidet der Einsatz von In-Memory-Technologie über die Zukunft ganzer Unternehmen. In einer Konkurrenzsituation, in der zwei Firmen ein Produkt anbieten, ist derjenige im Vorteil, der seine Kapazitäten und sein Risiko punktgenau kennt und in kürzester Zeit (im Idealfall sofort) ein maßgeschneidertes und zuverlässiges Angebot abgeben kann.

Ein potentielles In-Memory-Szenario könnte sich folgendermaßen darstellen:

Ein Kunde möchte ein Gut bei Ihnen kaufen. Sie nehmen seine Anfrage entgegen und erhalten binnen weniger Sekunden alle Parameter für Ihre Entscheidung zurück. In Echtzeit sehen Sie ihre globale ATP-Menge². Produzieren Sie alle 10 Sekunden ein Stück des Gutes und liefern Ihre Maschinen die Information zur Fertigstellung, so wird die gleiche Abfrage 10 Sekunden später

² ATP = Available-to-promise: zum Auftragszeitpunkt verbindlich zusagbarer Bestand

auch dieses zusätzliche Stück berücksichtigen. Wenn vor fünf Minuten eine für die Produktion wichtige Maschine ausgefallen ist, wird auch dieses in Ihrer ATP-Menge berücksichtigt. Auch auf der Kundenseite erhalten Sie alle aktuellen Informationen, sehen sein Kundenkonto mit der Einzahlung, die vor einer halben Stunde einging, ebenso wie das daraufhin aktualisierte Kundenranking. Ein Börsenportal und ein interner Informationsdienst melden, dass Ihr Kunde gestern den Zuschlag für einen Großauftrag von einem multinationalen Konzern erhalten hat. Sie sehen in Echtzeit wie sich der Verkauf auf Ihren Cash-flow auswirkt.

Die exakte Kenntnis der entscheidenden Einflussfaktoren in diesem Szenario erlaubt die Optimierung bestehender Geschäftsabläufe und die Erschließung neuer Potentiale. Unter dem Aspekt des Risikomanagements kommt In-Memory Technologie eine besondere Bedeutung zu: Echtzeit-Management kann Fehlentscheidungen nicht grundsätzlich verhindern, kann aber deren Häufigkeit reduzieren, da zumindest Fehler auf Grund von veralteten oder unvollständigen Informationen seltener werden.

SAP BW – Business Intelligence mit SAP HANA

Für viele Unternehmen ist Business Intelligence der Einstiegspunkt für In-Memory-Technologie. Im analytischen Umfeld kann SAP HANA seine Muskeln spielen lassen. Nicht zuletzt deswegen wurde SAP HANA bewusst zunächst als Appliance für das SAP Business Warehouse positioniert. SAP HANA ist zwar nicht grundsätzlich auf das Business Warehouse von SAP angewiesen, aber da die Verwendung von SAP HANA im Zusammenhang mit SAP BW wohl in den kommenden Jahren einen der häufigsten Anwendungsfälle darstellen dürfte, soll er hier zunächst betrachtet werden.

In den Anfängen der SAP kannte man die Trennung zwischen ERP und BW System nicht. Das Reporting fand im ERP System statt. Das hatte durchaus Vorzüge, denn die Daten mussten nicht erst aus dem ERP System in das BW System geladen werden. Mit steigenden Datenvolumina stellte man fest, dass dieses Nebeneinander nicht problemlos war. Es kam zu Sperrsituationen, wenn mehrere Benutzer gleichzeitig auf den

gleichen Datenbestand zugegriffen. Bei größeren Tabellen und komplexen Abfragen dauerten viele Analysen inakzeptabel lange und belasteten obendrein auch die Performance transaktionaler Abläufe. Aus dieser Not entstand die Idee die Daten aus dem ERP System in ein dediziertes System zu kopieren und dort die Analysen durchzuführen. Es entstand die Unterscheidung zwischen transaktionalem System (ERP) und analytischem System (BW). Wichtig ist an dieser Stelle festzustellen, dass dadurch viele Daten in beiden Systemen vorlagen (Redundanz).

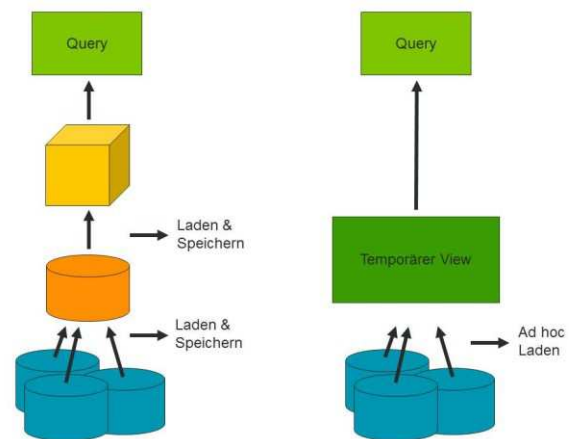
Auch die analytischen Systeme stießen durch das stetig steigende Datenvolumen bald an ihre Grenzen und man begann spezielle Objekte (z.B. InfoCubes) innerhalb des analytischen Systems zu entwickeln, deren einziger Zweck es ist, große Datenmengen so zu strukturieren, dass sie schneller analysiert werden können. Als auch dieses Vorgehen nicht mehr ausreichte entwickelte man höher spezialisierte Objekte (Aggregate), die nur noch einen Teilaspekt der Ursprungsdaten abbildeten. Den Vorteil der schnelleren Analyse können Aggregate jedoch nur ausspielen, wenn die Analyse sich auf genau die Daten im Aggregat bezieht. Bei der Modellierung von InfoCubes und Aggregaten muss folglich schon sehr genaue Kenntnis der Analyseanforderungen vorliegen. Diese Voraussetzung ist jedoch oftmals nicht gegeben. Ein Mangel an Flexibilität, häufige und aufwändige Änderungen und zusätzliche Redundanzen durch immer neue Objekte sind die Folge.

Der Aufwand für die Modellierung und Pflege der Objekte ist einer der zentralen Kostentreiber und damit ein wesentlicher Nachteil der Architektur klassischer BW Systeme. Dabei durchlaufen die Daten auf dem Weg von der Datenquelle bis in das Analyseobjekt (InfoCube) im Allgemeinen mehrere Schichten und werden dabei bereinigt, verfeinert, transformiert und für den Zweck der jeweiligen Analyse angepasst.

Diese Komplexität und die damit einhergehende Inflexibilität, machen Änderungen an derartigen Datenmodellen sehr schwierig und der vermeintlich „kleine“ Wunsch des Fachbereichs, in einer Analyse nur eine zusätzliche Kennzahl zu sehen, kann einen Modellierungsaufwand von mehreren Wochen bedeuten. Des Weiteren ist zu beachten,

dass derartige Architekturen nur unter großem Aufwand aktuelle Daten zur Verfügung stellen können. Generell müssen Daten auf ihrem Weg in die analytischen Objekte die verschiedenen Schichten in einer genau definierten Reihenfolge durchlaufen. Solche Stagingprozesse können, je nach Datenvolumen, mehrere Tage dauern. Zwar gibt es Analyseobjekte die Daten direkt aus der Quelle lesen können. Da sie jedoch nur kleine Datenvolumina verarbeiten können, werden sie nur ergänzend eingesetzt und bedeuten in jedem Fall einen zusätzlichen Modellierungs- und Wartungsaufwand.

Der ab 2005 von SAP zusätzlich eingeführte BW Accelerator (BWA) – eine Vorläufertechnologie von SAP HANA, ebenfalls basierend auf In-Memory-Technologie - macht Teile der Modellierung überflüssig, vornehmlich die Aggregate. Der BWA wirkt aber nur in den oberen Schichten des BW und bringt daher nur teilweise Einsparungen und Flexibilitätsgewinne.



Klassische DWH Architektur (links) versus SAP HANA Datenmodell (rechts)

Für Kunden mit einer SAP-Systemlandschaft mit einem SAP BW als zentralem Data Warehouse und Reportingwerkzeug lohnt sich also ein genauer Vergleich zwischen einem „Weiter so“ und dem Umstieg auf SAP HANA. Die Werkzeuge der SAP BW Workbench können weiterhin benutzt werden und auch bereits angelegte Objekte können auf SAP HANA weiterbetrieben werden, aber es ist wichtig zu sehen, dass neu entwickelte analytische Modelle wesentlich einfacher aufgebaut

werden, weitaus einfacher zu warten sind und was vielleicht einer der wichtigsten Aspekte ist, viel einfacher zu verändern sind. Wer demnach eine sinnvolle Kostenanalyse machen will, muss auch seine aktuellen Aufwendungen für Berater und Programmierer den Kosten für SAP HANA gegenüberstellen.

Neben dem geringeren Modellierungsaufwand im SAP HANA Umfeld soll abschließend noch ein Aspekt betrachtet werden, der sich direkt aus dem vereinfachten Datenmodell ableitet. Als transaktionales und analytisches System noch Eins waren, stellten die Mitarbeiter der Fachbereiche die Daten aus den verschiedenen Tabellen in einen sinnvollen Zusammenhang. Mit der Aufteilung in zwei Systeme und der Spezialisierung der analytischen Umgebung, mussten die Fachbereiche viele ihrer Kompetenzen an anderweitig spezialisierte Modellierer abgeben, die aufbauend auf den Vorgaben des Fachbereichs ein Datenmodell erstellten. Oft genug kam es dabei zu unbefriedigenden, undurchsichtigen und damit schwer wartbaren Lösungen, weil A nicht verstand, was B wirklich wollte. Die einfachen Datenstrukturen und die Fähigkeit große Datenbestände ad-hoc zu analysieren, bringen im SAP HANA Umfeld Modellierung und Analyse wieder näher an den Fachbereich. Auf Grund der enormen Leistungsfähigkeit von SAP HANA ist es möglich, dass Modellierer sehr allgemeine Datenmodelle zur Verfügung stellen und es dem Fachbereich überlassen wird, daraus die benötigten Informationen zu gewinnen.

Realtime Analytics auf operativen Daten

Wer sich heute für SAP HANA als Basis für ein Business Warehouse entscheidet, beschleunigt zunächst die analytische Seite. Das Design von SAP HANA erlaubt jedoch auch ein Miteinander von transaktionalen und analytischen Operationen. Seit Anfang 2013 steht auch die SAP ERP Suite powered by SAP HANA zur Verfügung (s.u.). Zusätzlich liefert die SAP den SAP HANA Analytics Foundation Browser aus, ein HTML5 basiertes Werkzeug, das es erlaubt SAP HANA Views zu durchsuchen, zu taggen und nach Excel zu exportieren.

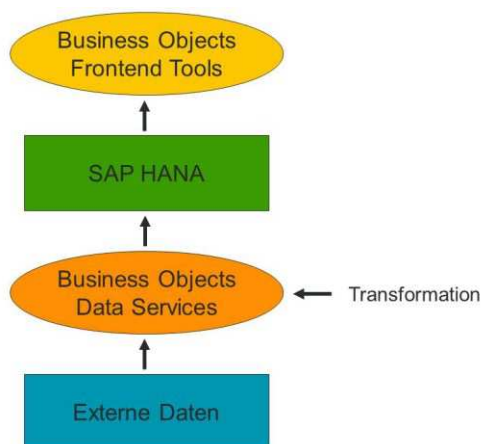
Die HANA-Strategie eines Unternehmens sollte diese Entwicklungen klar im Auge behalten, denn mit dieser Generation der Business Suite wird es – mindestens theoretisch - wieder möglich, dass operative Anwendung und Analyse auf der gleichen Plattform laufen. Zusätzliche Systeme, redundante Datenhaltung und das Kopieren von Daten zwischen transaktionalen und analytischen Umgebungen werden zukünftig in vielen Situationen wegfallen.

Bei der Erstellung einer Wirtschaftlichkeitsanalyse für ein Investment in SAP HANA sind die Kosten für Betrieb und Wartung mit „alter“ und „neuer“ Technologie die zentralen Aspekte. Analytische Anwendungen auf SAP HANA sind nicht mehr auf die typischen BW-Objekte wie InfoCubes und Aggregate angewiesen. Analysen im HANA-Umfeld basieren auf sogenannten flachen Tabellen. Das können im einfachsten Fall genau jene Tabellen sein, die im transaktionalen System mit Kundendaten, Bestelldaten etc. gefüllt werden. Zwar findet auch im analytischen HANA Umfeld noch eine Modellierung statt, diese ist jedoch im Vergleich zur komplexen Modellierung im BW sehr rudimentär und äußerst flexibel. Aufbauend auf diesem Datenmodell werden die Daten zum Zeitpunkt der Analyse ad hoc aus den flachen Tabellen gelesen. Damit entfällt nicht nur ein Großteil der Modellierungsarbeit. Die analysierten Daten sind auch immer so aktuell, wie sie in der zugrundeliegenden flachen Tabelle vorliegen. Die neue Generation der SAP Business Suite kann gemeinsam mit einem BW System auf einer SAP HANA Instanz betrieben werden, transaktionales und analytisches System greifen auf die gleiche Datenbasis zu und erlauben so Real-Time Analytics!

SAP HANA und Business Objects

Der Einsatz eines BW Systems, so wie man es aus dem SAP Umfeld kennt, ist im Zusammenhang mit SAP HANA nicht mehr zwingend notwendig. Ohnehin werden bei der Migration und dem Neuanlegen von Objekten im SAP BW Umfeld diese Objekte in SAP HANA konforme Strukturen überführt. Damit werden rechenintensive Operationen wie z.B. das Aktivieren von DSOs aus dem BW System auf die SAP HANA DB verlagert.

BusinessObjects Frontend Tools haben sich schon in der Vergangenheit wenig an der Idee des SAP BW orientiert. Sie waren vor der Übernahme von BusinessObjects durch SAP darauf hin entwickelt worden, Daten aus allen möglichen Datenquellen lesen zu können. Entsprechend ist es mit diesen Werkzeugen heute sehr einfach, die Daten aus SAP HANA direkt zu konsumieren. Dabei stehen alle bekannten Möglichkeiten offen (MDX, SQL, BICS, Universen). HANA bietet ein eigenes Sicherheitskonzept. Bei Verwendung der BO-Tools hat der Anwender daher die Wahl, Rollen- und Zugriffsrechte auf der HANA- oder BO-Ebene abzubilden.



Das Zusammenspiel von SAP HANA und SAP BusinessObjects

Von den BO Frontend Tools abschließend noch ein kurzer Blick auf die Eingangsschicht von SAP HANA, denn auch hier bietet BO Potential. ETL (Extraktion, Transformation, Loading) beschreibt die Prozesse, die bei der Integration von Daten aus Fremdsystemen in analytische Systeme Anwendung finden. Besteht z.B. die Notwendigkeit Daten auf ihrem Weg nach SAP HANA zu verändern oder sollen Daten aus externen Datenquellen in bestimmten Zeitabständen geladen werden, muss ein ETL Tool eingesetzt werden. Die BO DataServices erfüllen genau diesen Zweck und können diesbezüglich SAP BW ersetzen.

Die SAP Business Suite und SAP HANA

Seit Januar 2013 steht nun auch die *SAP Business Suite powered by SAP HANA* zur Verfügung und setzt damit neue Maßstäbe im Bereich Enterprise Resource Planning. Schon länger bietet die SAP sogenannte Rapid Deployment Solutions (RDS) an, von denen viele SAP HANA nutzen können. Im Zusammenhang mit SAP HANA werden diese Accelerators genannt, da sie Teilbereiche der Business Suite beschleunigen. RDS kapseln bestimmte Funktionen der Business Suite wie z.B. das Financial Reporting und werden als Komplettpakete von der SAP zu einem Fixpreis angeboten. Durch die Verlagerung rechenintensiver Operationen aus der Business Suite auf SAP HANA wird ein Geschwindigkeitsvorteil für die entsprechende Funktion erzielt.

Die RDS wurden von der SAP als Lösung für ein ansonsten oft ausuferndes und teures Projektwesen positioniert und haben im Bereich einer Standardimplementierung sicherlich ihre Berechtigung. Softwaretechnisch lassen sie sich zwischen der Business Suite und den By-Design Lösungen ansiedeln.

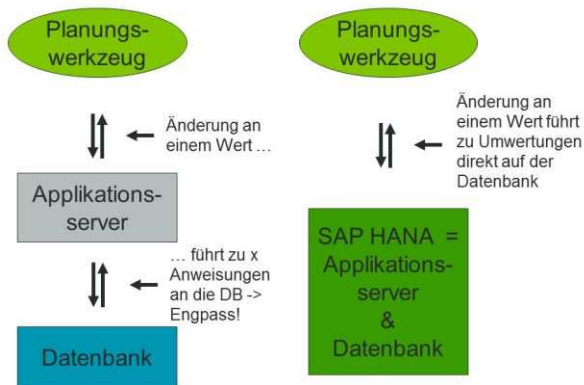
Mit dem Erscheinen der *SAP Business Suite powered by SAP HANA* werden diese Acceleratoren als eigenständige Lösung an Bedeutung verlieren. Eine Neuimplementierung oder ein Upgrade sollte jetzt auch immer die Möglichkeiten der *SAP Business Suite powered by SAP HANA* mit in Betracht ziehen. Der bereits erwähnte *SAP HANA Analytics Foundation Browser* bietet gute Möglichkeiten, Daten direkt im ERP System zu analysieren und für die weitere Bearbeitung im xml-Format zu exportieren. Und selbst fast schon vergessene Tools wie der Report Painter und der Report Writer, Urgesteine der ERP Lösung, erfahren mit der Business Suite powered by SAP HANA eine Renaissance.

Die *SAP Business Suite powered by SAP HANA* ist nicht nur für Kunden mit bereits existierender SAP-Installation interessant. Nun, da transaktionale Daten direkt in die HANA Datenbank geschrieben werden können, stehen diese dort dann auch sofort für die Analyse bereit und erlauben damit eine echte Real-Time-Analyse.

Planung mit SAP HANA

FSCM, APO, BPC, integrierte Planung... SAP hat eine lange Historie von Planungsanwendungen und verfügt über ein vielfältiges Spektrum an Tools für diesen Zweck. In kaum einem Bereich wird der Bedarf nach schierer Rechenleistung so offensichtlich wie in der Planung. Eine Einzelwertänderung in einer Planungsanwendung kann Rechenarbeit für mehrere Tage bedeuten und in der Vergangenheit war so mancher technische „Kunstgriff“ notwendig um die für komplexe Planungsszenarien notwendige Performanz zu erreichen.

SAP HANA verfügt über eine integrierte Planungseingine sowie planungsspezifische Funktionen. Die Planungsfunktionen können aktuell direkt über das HANA Studios modelliert werden. Die SAP arbeitet mit Hochdruck an der Bereitstellung eines neuen Planungswerkzeuges welches aller Voraussicht nach viele der aktuell verwendeten Tools ersetzen wird.



Planung klassisch (links) und mit SAP HANA (rechts)

Die enge Verknüpfung zwischen analytischen Funktionen und Planung macht die Integration beider Bereiche in einem System sinnvoll. Die SAP hat dieses bei der Entwicklung von SAP HANA berücksichtigt und die Grundlagen für ein erfolgreiches Miteinander beider Anwendungsbereiche geschaffen. Das Design von SAP HANA ist darauf ausgelegt, Daten auf granularer Ebene zu speichern und erst bei Anforderung durch ein analytisches Modell zu aggregieren. Diese Architektur

begünstigt Planungsanwendungen gleichermaßen.

Auch mit SAP HANA werden Planungsläufe noch Zeit in Anspruch nehmen, vor allem aufgrund der fachlich komplexen Planungsprozesse. Die Dauer dieser Prozesse wird sich jedoch durch die Verkürzung der technischen Prozessabläufe vermindern. Ein fehlerhafter Parameter bedeutet dann nicht mehr einen oder gar mehrere Tage Verzögerung, Produktionsausfall, höhere Zinsen...! Planung der Zukunft bedeutet Planung mit SAP HANA.

SAP HANA als Motor für Innovation

SAP HANA ist eine Appliance, also eine Kombination aus Hardware und Software.

Für Entwickler präsentiert sich SAP HANA als Datenbank mit einer SQL- und einer MDX-Schnittstelle sowie eingebauter Scripting-Funktionalität. SAP HANA ist in seiner Funktionalität eigenständig und in keiner Weise an andere SAP Produkte gebunden. Die Appliance kann also wie jedes „normale“ Datenbanksystem betrieben werden.

SQL Script, die Scripting Sprache in SAP HANA besteht aus einem Satz eingebauter Funktionen für wiederkehrende Datenbankoperationen, es besteht aber auch die Möglichkeit benutzerdefinierte Funktionen zu erstellen. Dadurch ist es möglich, komplexe Operationen in einer Funktion zu kapseln und nur die Schnittstelle nach außen zu publizieren. Ein besonderer Reiz benutzerdefinierte Funktionen liegt in der Möglichkeit, der benutzerdefinierten Funktion mitzugeben, in welcher Sprache die Funktion geschrieben ist. Auf diese Weise kann z.B. aktuell das gesamte Spektrum von R, einer im statistischen Umfeld äußerst populären Sprache, in den Ausführungsplan der HANA SQL Engine integriert werden.

Wenn von „Big Data“ gesprochen wird, fallen fast zwangsläufig auch die Begriffe „structured“ und „unstructured“ Data. „Structured Data“ bezeichnet dabei Daten die in strukturierter Form, also in Tabellen vorliegen. Der Informationsgehalt solcher Daten lässt sich im Allgemeinen gut erschließen da der Tabelleninhalt bereits mit einer gewissen Semantik verknüpft ist und Abfragesprachen wie SQL diese gut analysieren können.

„Unstructured Data“ bezeichnet Daten wie z.B. Fließtext. Besonderes Interesse erwecken in diesem Zusammenhang immer wieder die vielen Informationen, Blogs, etc. in Sozialen Medien. Diesbezüglich hat SAP einen Dinosaurier wiederbelebt, TREC. Beim – Text Retrieval and Extraction System der SAP handelt es sich um eine Suchmaschine die über eine sogenannte Fuzzy-Suche nach Ähnlichkeiten und Mustern suchen kann. Man spricht in diesem Zusammenhang auch von einer „unscharfen“ Suche, die mit normalen SQL Abfragen nicht zu realisieren ist. Während andere Suchmaschinen ihren Zenit erlebten verrichtete TREC, ganz ungewohnt für einen Fleischfresser, geruhsam und bescheiden seine Dienste im Bereich SAP Enterprise Search.

Die Logik, die TREC zugrunde liegt, steht nun auch in SAP HANA zur Verfügung. Damit kann jede dort abgelegte Information über eine Fuzzy-Suche analysiert werden. Die geballte Kraft von SAP HANA macht nun aus dem Dinosaurier das, was er schon immer war: eine Textfressmaschine, die in rasanter Geschwindigkeit Texte verschlingt. Wer den Dinosaurier zu zähmen weiß und schaut was rauskommt, der erkennt vielleicht den Goldesel!

Eine Vielzahl von Anwendungen wird durch die enorme Leistungsfähigkeit von SAP HANA erst möglich. Genomanalysen beispielsweise und andere äußerst rechenintensive Anwendungen sind bereits erfolgreich demonstriert worden: www.experiencesaphana.com bietet einen guten Einstiegspunkt für weitere Beispiele und zusätzliche Informationen.

SAP HANA ist noch jung und viele Anwendungsmöglichkeiten müssen erst noch erdacht werden. Der Vergleich mit einer x-box hat durchaus seine Berechtigung, denn so wie für diese immer neue Spiele entstehen, könnten und werden für SAP HANA immer neue Anwendungen entstehen.

Technische Grundlagen und Sicherheitsaspekte der „In-Memory“ Technologie

Die Verwendung von 32-bit Betriebssystemen begrenzte lange Zeit den maximal verwendbaren Arbeitsspeicher auf 4G-Byte. Mit der Einführung von 64-bit Betriebssystemen wurde diese Grenze aufgehoben. Derartige Systeme waren jedoch

lange Zeit Spezialanwendungen vorbehalten. Durch sinkende Preise für Arbeitsspeicher und die zunehmende Verbreitung von 64-bit Betriebssystemen entwickelte sich ein Umfeld, das die Nutzung von Rechnersystemen, die nur noch aus Arbeitsspeicher bestehen, lukrativ machte. Der Vorteil einer solchen Architektur ist, dass alle Daten und Programme immer im schnellen Arbeitsspeicher vorhanden sind. Es müssen keine Daten mehr von einer langsamen Festplatte geladen oder auf diese ausgelagert werden.

Der BW Accelerator war der erste Schritt der SAP in Richtung In-Memory-Technologie. Mit seinem 64-bit Betriebssystem war der BWA in der Lage, große Mengen Arbeitsspeicher zu verwalten. Dadurch konnten Datenmodelle (InfoCubes) als Ganzes im Arbeitsspeicher gehalten und so die Antwortzeiten für Analysen auf diesen Datenbeständen drastisch reduziert werden. Viele der Technologien aus dem BWA finden sich heute in SAP HANA wieder, aber während der BWA als reines Zusatzprodukt zu SAP BW konzipiert wurde und nur mit den dort verwendeten InfoCubes einen Geschwindigkeitsvorteil erreichen kann, ist SAP HANA eine vollkommen eigenständige Lösung.

Name	PLZ	Umsatz
Kunde A	1234	250
Kunde B	2345	600
Kunde C	1234	150
Kunde D	1234	750

Zeile 1	Kunde A	Spalte 1	Kunde A
	1234		Kunde B
	250		Kunde C
Zeile 2	Kunde B	Spalte 2	Kunde D
	2345		1234
	600		2345
Zeile 3	Kunde C	Spalte 3	Kunde A
	1234		1234
	150		250
Zeile 4	Kunde D		600
	1234		150
	750		750

Speichernutzung bei Zeilen- (links) und Spaltenorientierung (rechts)

Bereits mit dem BW Accelerator führte die SAP neben dem klassischen zeilenbasierten Datenmodell ein spaltenbasiertes Datenmodell ein. Das zeilenorientierte Datenmodell betrachtet die einzelnen Datensätze als in Zeilen abgelegt wäh-



rend die Spalten die Feldbezeichnung enthalten. Eine Kundendatei mit 1 Million Kunden enthält z.B. zu jedem Kunden die Kundennummer, den Namen, ein Feld „Postleitzahl“ und „Telefonnummer“. In der Spalte Postleitzahl sind häufig die gleichen Einträge zu finden und das Feld Telefonnummer bleibt oft leer. Das zeilenbasiertes Datenmodell kann solche Strukturen nur sehr bedingt optimieren. Das spaltenbasierte Datenmodell betrachtet die einzelnen Spalten und kann so z.B. in der Spalte „Postleitzahl“ eine verbesserte Umkodierung und Komprimierung erzielen, Kennzahlen in angrenzenden Speicherbereichen sind sehr leicht zu aggregieren.

Die spaltenbasierte Datenablage unterstützt moderne Prozessor- und Speicherarchitekturen und beschleunigt dadurch SAP HANA zusätzlich. In der Diskussion um die geeignete Hardware wird oft der Vorwurf erhoben, die SAP zwingt Kunden, eine bestimmte Hardware zu kaufen. Richtig ist, dass nur zertifizierte Hardwarepartner die fertig konfigurierte Appliance ausliefern können. Eine Installation auf beliebiger Hardware ist nicht möglich. Dieses hat jedoch durchaus seine Gründe. Wesentlich zu beachten ist, dass SAP HANA aktuell nur mit Chips die der Nehalem bzw. Westmere Architektur entsprechen, seine volle Leistungsfähigkeit entfalten kann.

Letztlich noch ein kurzer Blick auf zwei wichtige Themen: Transaktionssicherheit und Ausfallschutz. Transaktionssicherheit war einer der wesentlichen Erfolge für die Software aus dem Hause SAP und so gilt auch im SAP HANA Umfeld die bekannte Devise: Alles oder Nichts! HANA verfügt über die aus dem SAP-Umfeld bewährten Mechanismen und eine Transaktion wird entweder komplett abgeschlossen oder komplett zurückgerollt.

Was den Ausfallschutz betrifft, so kommen hier - und nur hier - Festplatten zum Einsatz. Wie bei

jeder Datenbank, so muss auch bei der HANA Datenbank kontinuierlich ein Backup gemacht werden, um einen Ausfall der Datenbank abzufangen. Eine HANA Appliance wird bereits vorkonfiguriert mit der entsprechenden Hardware ausgeliefert. Auch was die Mechanismen wie Logpoints und Savepoints betrifft, unterscheidet sich SAP HANA nicht von einer herkömmlichen Datenbank.

Der Autor

Markus Alterauge ist Berater im Bereich SAP der mayato GmbH und Technologieexperte für SAP BW und HANA.

Markus Alterauge | mayato GmbH
markus.alterauge@mayato.com

Über mayato

mayato ist als Analytisten- und Beraterhaus, spezialisiert auf Business Intelligence. Von Niederlassungen in Berlin, Bielefeld und Heidelberg aus arbeitet ein Team von erfahrenen IT- und BI-Architekten, Statistikern, Analytisten sowie fachlichen Experten für spezielle Themen wie Betrugs-erkennung, Data Mining und Analytisches CRM. Zu den Kunden von mayato zählen namhafte Unternehmen aus unterschiedlichen Branchen. Als Partner mehrerer Softwareanbieter ist mayato grundsätzlich der Neutralität und in erster Linie der Qualität seiner eigenen Dienstleistungen verpflichtet. Nähere Infos unter: www.mayato.com.

mayato GmbH

Am Borsigturm 9
D-13507 Berlin
T +49 30.4174.8657
info@mayato.com
www.mayato.com