

EFFIZIENTE NEARLINE-DATENHALTUNG ALS BAUSTEIN GROSSER SAP BW-IMPLEMENTIERUNGEN

Daten suchen & finden

Unternehmen brauchen heute mehr denn je aktuelle Informationen und detaillierte Einblicke in ihre Geschäftsaktivitäten. Um diese Informationen zu erhalten, setzen sie weltweit zunehmend auf Data Warehouses. Dort werden Daten sinnvoll zusammengeführt, nützliche Informationen aufbereitet und auf deren Basis fundierte Entscheidungen getroffen.

Von Roland Markowski, Dr. Michael Hahne

Dabei handelt es sich allerdings in den seltensten Fällen um statische Anwendungen. Im Gegenteil, aufgrund stetig steigenden Wettbewerbsdrucks, wiederkehrender Gesetzesänderungen und sich verändernder Anwenderbedürfnisse muss sich ein Data Warehouse ständig an folgende Erfordernisse anpassen: Verfügbarkeit weltweit, 24 x 7, ad hoc; Einsatzbereiche, unternehmensweit und -übergreifend; Aktualität, zeitnahe Informationen; Granularität, Zugriff auf detaillierte Daten und Zeiträume, Zugriff auf historische Daten.

Je nach Unternehmensgröße, Branchenzugehörigkeit und Anspruch an Datengranularität und -historie müssen Unternehmen mit zum Teil sehr hohen Wachstumsraten bei ihren Datenmengen umgehen. Dies stellt auf der einen Seite eine technische Herausforderung dar und ist andererseits mit nicht unerheblichen Betriebskosten verbunden. Hierzu Gartner-Analyst Mark Beyer: „In Data Warehouses ohne wirksame Strategie zur Archivierung und Handhabung von Daten, die nicht mehr aktiv genutzt werden, wächst die Datenmenge pro Jahr um mehr als 100 Prozent - und zwar noch bevor neue gesetzliche Vorgaben bei der Datenhaltung berücksichtigt werden können. Vor diesem Hintergrund müssen Data-Warehouse-Manager ihre IT-Lösungen kreativ einsetzen, um ein Gleichgewicht zwischen Wachstum und Systemleistung zu finden.“

Für Data-Warehouse-Manager, die das Business Information Warehouse (SAP BI) einsetzen, bedingt die Erfüllung der genannten Ansprüche

der Benutzer oft den Umbau der Data-Mart-getriebenen SAP BI-Architektur hin zu einer Architektur mit deutlich höherem Anteil an detaillierten Daten in ODS-Objekten das Vorhalten von Daten aus noch umfangreicheren Zeiträumen. Als Ergänzung dieses Umbaus des SAP BI kann sich zudem die Einführung einer Nearline-Komponente für das Gleichgewicht zwischen Datenwachstum und Systemleistung als sinnvoll und Kosten senkend erweisen. Dies bedeutet, dass z.B. ältere oder selten genutzte InfoCubes oder ODS-Objekte auf andere Speichermedien ausgelagert werden und trotzdem für die Benutzer im Zugriff bleiben. Dazu ist SAP Software-Partnerschaften eingegangen, u.a. auch mit SAND Technology auf der Basis des SAND Searchable Archive.

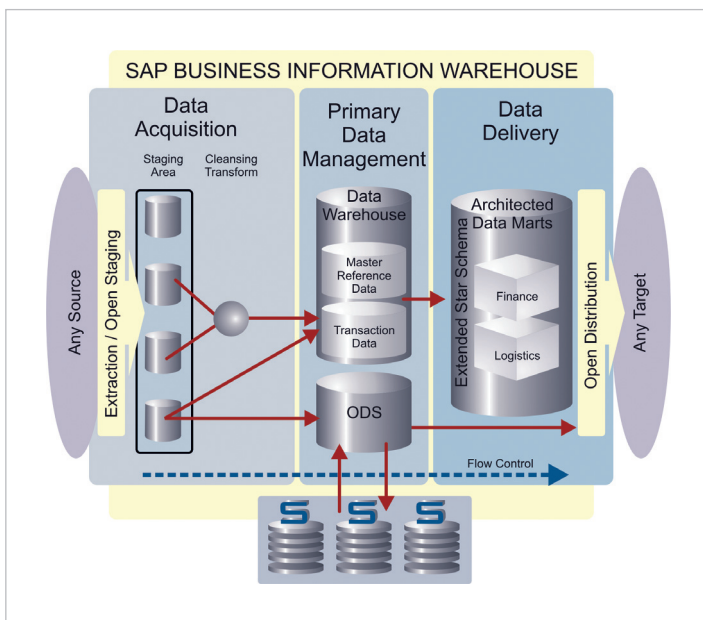
Das SAND Searchable Archive ist eine von der Hardware unabhängige Softwarelösung zur effizienten und Platz sparenden Speicherung strukturierter Daten. Mittels eines parallelisierbaren Ladeprozesses werden zu speichernde Tabellen sowie deren Spalten in der Regel > 85 % komprimiert und, gesteuert über die eigene Metadatenverwaltung, auf den zugewiesenen Speicherpartitionen in einem speziellen Format (SAND Compacted Table) abgelegt.

Damit ist gleichzeitig die Aufbereitung aller Tabellen und ausnahmslos aller Spalten für Abfragen ohne nachfolgende spezielle Indexierung verbunden. Das SAND Searchable Archive präsentiert sich nach dem Laden der Daten für den Benutzer wie eine „herkömmliche“ Datenbank, d.h. Abfragen erfolgen dann ohne vorherige De-

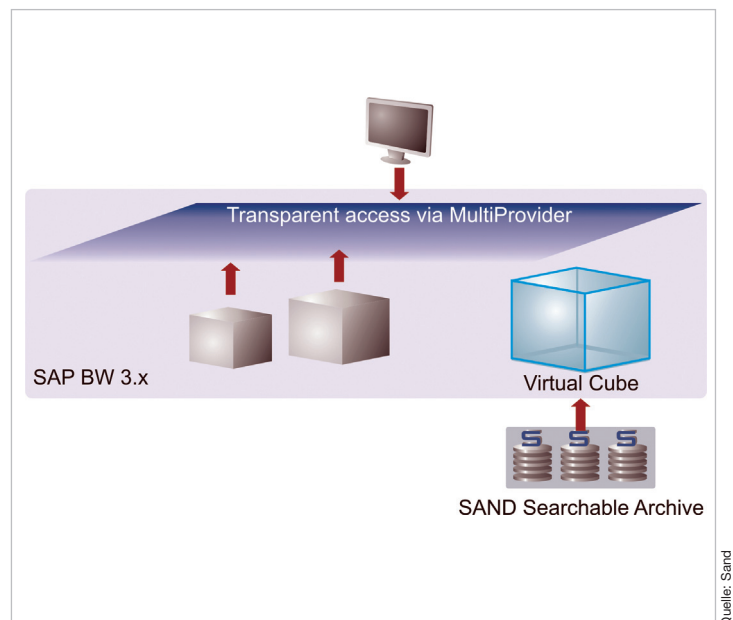
komprimierung – bei entsprechender Hardwarekonfiguration parallelisierbar – mittels SQL direkt auf den Tabellen sowie deren Spalten im komprimierten Zustand. Ein mit ausgelieferter Workflow-Manager sorgt bei Bedarf für die Integration in die betrieblichen (Batch-) Prozesse.

Auf diese Weise lassen sich die im BI gespeicherten Info-Provider hocheffizient (i.d.R. mehr als 85 Prozent) komprimieren, ohne dadurch den Zugriff auf die Info-Provider für den SAP BI-Prozessablauf oder die direkte Analyse zu verlieren. Administratoren von SAP BW 3.x können das SAND Searchable Archive über die zentrale, leicht bedienbare grafische Oberfläche steuern. Über das Zertifizierungsprogramm von SAP wird dies auch zu einer integrierten Nearline-Lösung in SAP NetWeaver 2004s werden. Die Integration des SAND Searchable Archive versetzt nun Unternehmen mit SAP BI-Anwendungen in die Lage, diese bezüglich Detaillierungsgrad und Zeithorizont effizient zu erweitern. Gleichzeitig können entsprechend „große“ SAP BI-Anwendungen bezüglich der Datenhaltung „schlanker“ gestaltet werden.

Neben dem rasanten Datenwachstum sieht sich ein Data-Warehouse-Manager in der Praxis häufig – über die Jahre gewachsen und aufgrund der unterschiedlichen Bedürfnisse der verschiedenen Abteilungen und Entscheidungsebenen – mit einer Vielzahl von unterschiedlichen technischen Lösungen konfrontiert. Dabei erweist es sich als aufwändig und kostenintensiv, diese zu erweitern oder fortzuentwickeln. Die Folgen sind unzuverlässige Geschäftsinformationen, verzögerte Entscheidungen und damit einhergehend die mangelhafte Umsetzung von Geschäftszielen. Weltweit operierende Unternehmen haben dies als Wettbewerbsnachteil erkannt und führen deshalb unternehmensweite und umfassende Enterprise Data Warehouses auf Basis eines SAP Enterprise Data Warehouse mit folgenden Schichten (layer) ein: Staging layer; Data Warehouse layer; Data Mart layer; und Operational Data Store layer. Die erste Schicht deckt den Bereich des Staging ab und dient der Ablage der Quelldatenextrakte auf Belegebene (row level). Ziel der Data-Warehouse-Schicht ist der Aufbau des integrierten Datenmaterials



Der SAP Datenbestand ist in jedem Fall das Rückgrat des Unternehmens: Gesucht und gefunden wird mit anderen Werkzeugen.



Für den Endanwender ist der transparente Zugriff auf heterogene Daten der entscheidende Erfolgsfaktor bei Business Intelligence

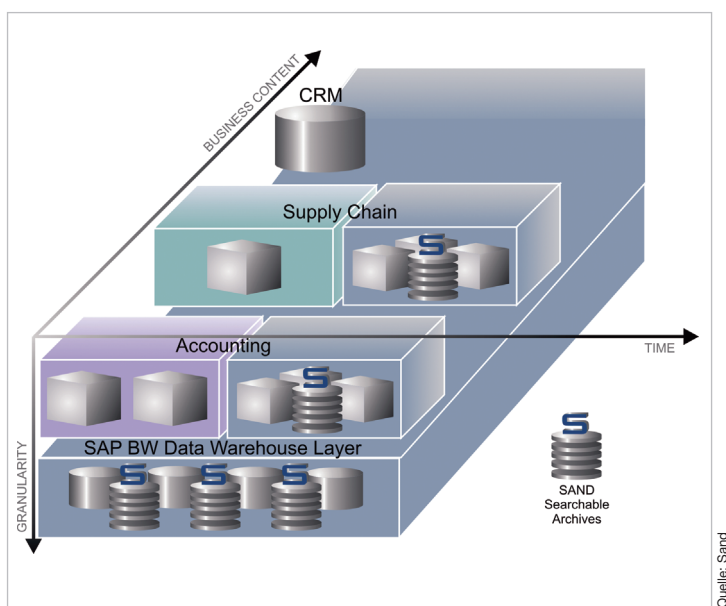
auf granularer Ebene. Diese stellt damit die Historie zur Verfügung und ist gleichermaßen unabhängig von der (betriebswirtschaftlichen) Anwendung. Die Data-Mart-Schicht bietet nun ihrerseits genau diese Sicht eines spezifischen Anwendungsbereichs und stellt dafür aggregierte, integrierte Daten zur Verfügung. Ebenfalls anwendungsspezifisch ist der Operational Data Store, in diesem sind die Daten jedoch granular abgelegt. Diese Schicht bedient die Anforderungen an ein Realtime-nahes Data Warehousing, da hier kurze Aktualisierungszeiten möglich sind.

Zu beachten ist, dass diese Schichten-

Architektur rein konzeptionell ist, die physische Ebene und die Strukturierung verteilter Systeme zu einer gesamten Data-Warehouse-Landschaft ist eine andere Fragestellung, die unter dem Begriff der Topologie getrennt zu diskutieren ist. Damit ist der SAP BW Data Warehouse Layer die einzige Quelle für die Data Marts des SAP BW und ermöglicht die umfassendere und schnellere Erfüllung neuer, auch ausgefallener analytischer Anforderungen, da neue Data Marts mit Historie und ohne erneute Extraktion aus den operativen Systemen zeitnah aufgebaut werden. Gleichzeitig werden redundante Staging-Prozesse vermieden sowie Extraktion und Integration zentral verwaltet.

Effiziente Verwaltung von großen Datenmengen im SAP BW Data Warehouse Layer. Das Datenvolumen hat sich als limitierender Faktor für Antwortzeitverhalten sowie Stabilität und Geschwindigkeit dieser ETL-Prozesse herauskristallisiert. Dies führt teilweise zu steigenden Ladezeiten sowie zu Fehleranfälligkeiten. Gleichzeitig ist die Umsetzung weiter gehender Anforderungen gerade aufgrund dieser Grenzen nicht möglich. Daher ist es sinnvoll, den SAP BW Data Warehouse Layer zur Verwaltung sehr großer Datenmengen mit dem SAND Searchable Archive zu ergänzen. Dank des hohen Datenkomprimierungsgrads können nun Daten tieferer Daten-ebenen, granulare „gealterte“ Detaildaten, nicht regelmäßig benutzte Daten, „Legacy“-Daten und eventuell für Analyse Zwecke notwendige Daten für die rasche Adaption neuer Analyseanforderungen Speicherplatz sparend vorgehalten werden. Diese stehen dann ohne Einschränkung für direkte Abfragen oder für den Aufbau daraus abzuleitender Info-Provider zur Verfügung.

Generell gilt: Neben dem Aufwand von einem Tag für die Einrichtung des SAND Searchable Archive selbst und für die Konnektivität zum SAP BW entsteht ein Umstellungsaufwand, der von der Anzahl der betroffenen Info-Provider abhängt. Im Falle eines sauberen Multi-Provider-Konzepts besteht die Umstellung nur aus dem Splitten der Info-Provider und dem zusätzlichen Einhängen des virtuellen (Archiv-)Cubes in die Multi-Provider. Findet auf dem zu splittenden Info-Cube Reporting statt, so muss für eine integrierte trans-



Business Intelligence ist heute mehr als ein Data Warehouse mit Drill-down-Tools, hinzu kommt CRM, Portale, SCM, etc.

Quelle: Sand

parente Sicht im Reporting, also für Analysen auf den physisch persistenten Info-Providern, zusammen mit den Nearline-Cubes aus dem SAND Searchable Archive noch die Ebene des Multi-Providers eingefügt und das Reporting auf diese Ebene migriert werden.

Die Integration des SAND Searchable Archive ermöglicht das effiziente Verhalten von bisher nicht gekannten Datenmengen. Unternehmen mit SAP BI-Anwendungen erzielen damit folgende Vorteile:

- Höhere Akzeptanz beim Management und den Benutzern.
- Umfassendere und schnellere Adaption neuer, auch ausgefallener analytischer Anforderungen.
- Schnellere Reaktion auf Anforderungen der Benutzer und des Managements.
- Höherer Detaillierungsgrad der Daten inkl. einer zeitlich erweiterten Abdeckung.
- Stabilisierung des SAP BI durch schnelleres und einfacheres Software- und Release-Management sowie reduzierte Backup- und Recovery-Zeiten.
- Senkung der Gesamtkosten durch Reduktion der Hardware-Kosten für Platten, Hauptspeicher und CPU sowie der Kosten für die Systemadministration bzw. bessere Ausnutzung vorhandener Hardware-Kapazitäten.
- Bessere Performance durch kürzere Antwortzeiten im Dialog (je nach Architektur) und Beschleunigung von Ladeprozessen.
- Senkung der „Reporting-Aktivierungszeit“ der nicht-transaktionalen ODS-Objekte durch direkten Zugriff auf ODS-Objekte im SAND Searchable Archive.
- Vereinfachter Einstieg in das Enterprise Data Warehousing mit SAP BI sowie in die Standardisierung unternehmensweiter BI-Lösungen anderer Hersteller.

Quelle: www.sand.com

SAND Technology

SAND Technology liefert Unternehmen aller Branchen Software für effizientes analytisches Datenmanagement. Der geschäftliche Nutzen von Informationen und Daten ändert sich mit der Zeit. Dennoch werden häufig aktuelle und „veraltete“ Daten mit konventionellen Mitteln unter erheblichen finanziellen Anstrengungen verwaltet als seien sie gleichwertig und für Analysen in einem Data Warehouse bereitgestellt. SAND Technology geht hier neue Wege – und stimmt das analytische Datenmanagement auf den jeweiligen geschäftlichen Nutzen der Informationen ab: wichtige zu analysierende Daten erhalten hohe Performance und Speicherressourcen, wenig genutzte Daten dagegen werden hoch komprimiert und damit kostengünstig für die Analyse bereitgestellt.

Das Unternehmen unterhält Partnerschaften mit Atos Origin, Business Objects, CSC, cundus AG, HP, Hyperion, Mummert Consulting, SAP und SUN. Neben Deutschland ist SAND auch in den USA, Kanada und England vertreten.

cundus AG

cundus ist ein auf Business Intelligence spezialisiertes IT-Dienstleistungsunternehmen mit betriebswirtschaftlicher Beratungskompetenz. Das in Duisburg ansässige Unternehmen ist Experte für die Bereitstellung integrierter Managementinformationssysteme zur Aufbereitung, Analyse und Darstellung geschäftlich relevanter Daten. Das betriebswirtschaftliche Know-how kommt insbesondere beim Aufbau von Unternehmensberichtssystemen, Planungsanwendungen sowie Lösungen für Konsolidierung und Corporate Performance Management zum Tragen. cundus bietet Unterstützung bei der Auswertung und Integration analytischer Informationen aus unterschiedlichen Data-Warehouse-Systemen, insbesondere aus dem SAP Business Information Warehouse. Zu seinen Kunden zählen unter anderen Bayer, Henkel, Haniel und Roche. Die Verbindung von Business und IT ist die Grundlage der Tätigkeit und zugleich die Stärke von cundus.

cundus ist Servicepartner der SAP AG und Certified Partner der Microsoft Corporation. Außerdem unterhält das Unternehmen Kooperationen mit der SAP SI AG, MIS AG und arcplan AG.

Die Autoren:

Roland Markowski, Geschäftsführer SAND Technology Deutschland GmbH
Dr. Michael Hahne, Chief Technology Officer der cundus AG