

Anlage 1 zur SAP BW DV

allgemeine technische Beschreibung des SAP BW-Systems/Begrifflichkeiten

Das SAP NetWeaver Business Warehouse (SAP BW) ist eine Lösung des Softwareherstellers SAP und stellt eine Umsetzung des Konzepts des Data Warehouse dar. Im KIT wird derzeit die Version (Release) BW 7.3 eingesetzt.

Zur Definition eines Datenmodells im BW benötigt es einige Objekte, die im Folgenden vorgestellt und in Tab. 1 dargestellt werden.

I. Infoobjekte

Das InfoObjekt ist die Basis für die Definition aller Datenmodelle im BW und definiert ein Feld. Es gibt vier verschiedene Arten von InfoObjekten: Merkmale, Kennzahlen, Einheiten und Zeiten

1) Merkmale

Ein InfoObjekt oder auch Merkmal kann Attribute besitzen, welche wiederum selbst als Merkmal definiert sein müssen. Um die Merkmale später in einem Cube auch sinnvoll nutzen zu können, muss das Merkmal mit den entsprechenden Stammdaten „gefüllt“ werden, d.h. sie müssen aus dem Quellsystem in das Merkmal geladen und dort mit einer SID (Surrogate ID oder Stammdaten-identifikation) als Schlüssel gespeichert werden.

Die Attribute eines Merkmals dienen in erster Linie zur Anzeige. D.h. bei einer Abfrage können weitere Informationen über das Merkmal dargestellt werden. Möchte man das Abfrageergebnis im Voraus einschränken, kann man die Attribute zum Selektieren und Navigieren nutzen. Dafür muss das entsprechende Attribut jedoch als Navigationsattribut deklariert werden.

2) Kennzahlen

Die Kennzahlen enthalten im Gegensatz zum Merkmal Bewegungsdaten, also Daten die im täglichen Betrieb anfallen und sind für Berechnungen geeignet. Die Kennzahlen können eine Einheit besitzen.

3) Einheiten

Es handelt sich hierbei um eine Währung oder eine Mengeneinheit.

4) Zeiten

Werden immer dann eingesetzt, wenn ein Feld ein Datum und/oder eine Uhrzeit als Inhalt hat. Durch die Deklaration als Zeitmerkmal findet so automatisch die Konvertierung statt, bspw. zwischen verschiedenen Formaten.

II. Weitere Objekte im Zusammenhang mit der Datenhaltung bzw. Datenbeschaffung sind: DataStore-Objekte, InfoCubes, Multiprovider, DataSource:

1) Infoprovider: DataStore-Objekt (DSO),

Beim DSO handelt es sich um eine Tabellenstruktur, deren Aufgabe es ist, die Daten aus einem oder mehreren Quellsystemen zu speichern. Das DSO dient somit zum einen als Datenquelle für weitere Datenziele, z.B. InfoCubes, kann aber auch direkt zu Analysezielen genutzt werden. Das DSO besteht aus Schlüssel- und Datenfeldern, die jeweils durch InfoObjekte festgelegt werden. Es können maximal sechzehn Schlüsselfelder definiert werden, eines ist jedoch vorausgesetzt. Die hierfür ausgewählten Merkmale dienen als Primärschlüssel der Tabelle. In den Datenfeldern folgen dann alle anderen Merkmale, Kennzahlen, Einheiten und Zeiten, die später in den InfoCubes oder direkt bei der Analyse benötigt werden.

Abb. 1 zeigt den Aufbau eines DSO. Neben Infos und technischen Einstellungen zum DSO sind auch die zuvor genannten Schlüssel- und Datenfelder zu sehen. Besitzt ein Merkmal ein Navigationsattribut, kann es hier aktiviert werden. Wird es nicht aktiviert, kann es später bei der Analyse nicht zur Selektion oder Navigation genutzt werden.

DataStore-Objekt	Techn. Name / W...	F...
DSO Aufträge Kosten	ZPDPM001	
Info zum Objekt		
• Version	◇ in Bearbeitung	
• Sicherung	⊕ gesichert	
• Überarb. Version	≡ aktive Version	
• Objektstatus	⊞ aktiv, ausführbar	
▶ Einstellungen		
▶ Schlüsselfelder		
▶ Datenfelder		
▶ Navigationsattribute		
▶ Indizes		

Abb. 1 - Aufbau DSO (SAP System KIT)

2) Infoprovider: InfoCube

Das wichtigste Objekt ist der InfoCube. Er repräsentiert das multidimensionale Datenmodell im SAP BW. Es können Dimensionen angelegt und Kennzahlen bestimmt werden. Die Modellierung hierfür erfolgt durch das erweiterte Stern-Schema. Ein InfoCube kann insgesamt sechzehn Dimensionen besitzen, wobei dreizehn individuell definierbar sind. Die drei vorgelegten Dimensionen sind Paket, Zeit und Einheit. Die Paket-Dimension beinhaltet die Request-ID, „eine technische Detaillierungsstufe, die den Ladevorgang beschreibt, über den die Daten in einen InfoCube gelangt sind.“ Die Änderungs-ID und der Satztyp sind weitere technische Merkmale und haben mit dem vorher erstellten Datenmodell nichts zu tun. In der Zeit-Dimension werden alle Zeit-Merkmale definiert, genauso die Einheiten in der Einheit-Dimension.

Abb. 2 zeigt den Aufbau eines InfoCubes mit den vordefinierten Dimensionen, bevor weitere Modellierungen vorgenommen wurden.

InfoCube	Techn. Name / W...	F...
▼ TEST	ZCUBE_T	
▼ Info zum Objekt		
• Version	◇ neu	
• Sicherung	☐ nicht gesichert	
• Objektstatus	⊙ inaktiv, nicht au...	
▶ Einstellungen		
▼ Dimensionen		
▼ Datenpaket	ZCUBE_TP	
• Änderungslauf ID	0CHNGID	
• Satztyp	0RECORDTP	
• Request ID	0REQUID	
▼ Zeit	ZCUBE_TT	
▼ Einheit	ZCUBE_TU	
▶ Dimension 1	ZCUBE_T1	
▶ Navigationsattribute		
▶ Kennzahlen		

Abb. 2 - InfoCube Aufbau (SAP System KIT)

3) Infoprovider: Multiprovider

Der MultiProvider dient dazu, Daten aus mehreren InfoProvidern zusammenzuführen, um diese dann für Analysen/Queries zu nutzen. Der Aufbau eines MultiProviders ähnelt dem eines InfoCubes. Er besteht aus Kennzahlen und Dimensionen. Die Definition der Dimensionen dient jedoch nur zur besseren Nutzbarkeit für Analysezwecke und repräsentiert nicht die physische Speicherung der Daten. Neben den Dimensionen und Kennzahlen müssen noch die Felder festgelegt werden, welche im MultiProvider angezeigt werden sollen und welche Felder aus den verschiedenen InfoProvidern einander entsprechen, um die Daten richtig zusammenführen zu können.










4) Data Source

Das Objekt zur Datenbeschaffung im BW ist die DataSource. Sie beschreibt eine Menge von Stamm- oder Bewegungsdaten aus einem Quellsystem. Es handelt sich hierbei um eine Struktur, die im Quellsystem angelegt und dann in das BW kopiert wird. Diese Struktur enthält alle nötigen Daten als Felder und wird zum Laden der Daten aus dem Quellsystem genutzt. DataSources können zum einen zum Laden von Daten aus anderen SAP Systemen genutzt werden, zum anderen aber auch für andere Quellen, z.B. eine CSV-Datei. Je nachdem wofür sie genutzt wird, muss der Typ der DataSource festgelegt werden.

5) InfoSource

Die InfoSource dient zum Zusammenführen von mehreren Transformationen. Sollen z.B. aus mehreren DSOs die Daten in mehrere InfoCubes geladen werden, fasst die InfoSource die Transformationen zusammen und „verteilt“ sie dann wieder auf die InfoCubes.

Tab. 1 - Wichtige Objekte im SAP BW (eigene Tabelle)

Merkmal		DataStore-Objekt	
Kennzahl		InfoCube	
Zeit		MultiProvider	
Einheit		DataSource	
		InfoSource/ InfoSource 3.x	

III. Architektur/Datenmodellierung

Die Datenmodellierung im SAP BW erfolgt mittels der Layered Scalable Architecture (LSA). Sie besteht dabei aus sieben bzw. acht Schichten (Abb. 3)..

1. Data Acquisition Layer (Datenübernahmeschicht)
2. Quality and Harmonization Layer (Qualitäts- und Harmonisierungsschicht)
3. Data Propagation Layer (Datenbereitstellungsschicht)
4. Business Transformation Layer (Geschäftslogik-Transformationsschicht)
5. Reporting Layer (Berichts- und Analyseschicht)
6. Corporate Memory (Unternehmensgedächtnis)
7. Operational Data Store (Operationale Datenversorgung)
8. Virtualization Layer

Der Virtualization Layer ist in Abb. 3 nicht zu finden. In dieser Schicht werden im SAP BW die MultiProvider eingesetzt. In dieser Schicht werden keine Daten gespeichert, sondern lediglich die Sicht auf die Daten in der Reporting Schicht, je nach Anforderung der Abfrage, verändert.

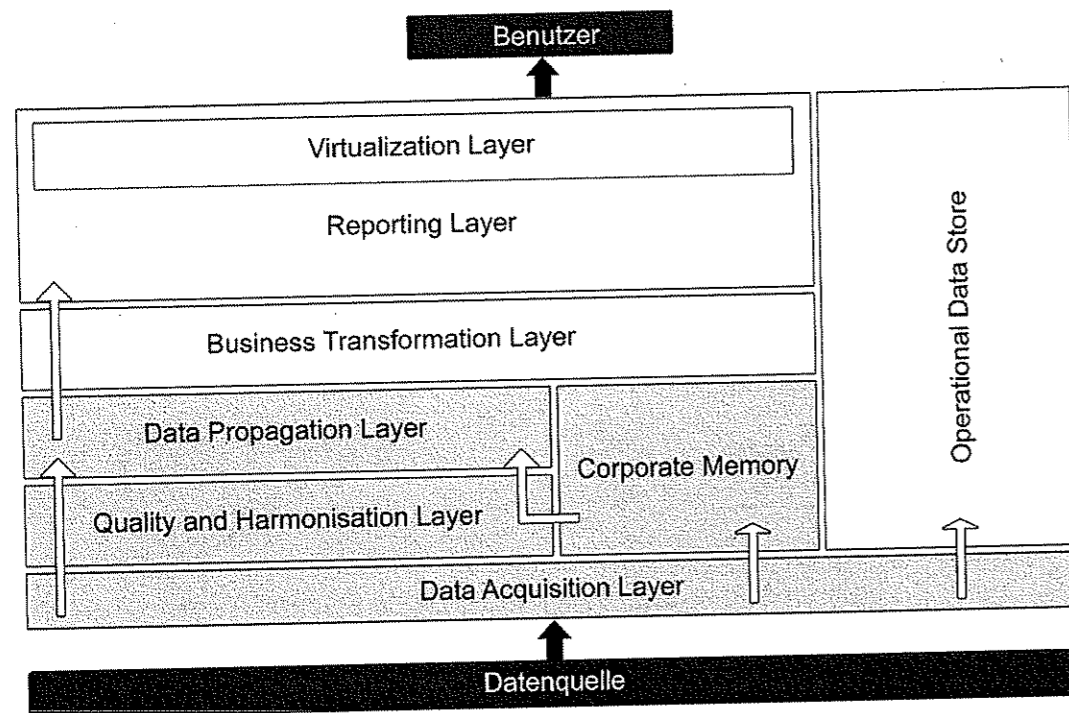
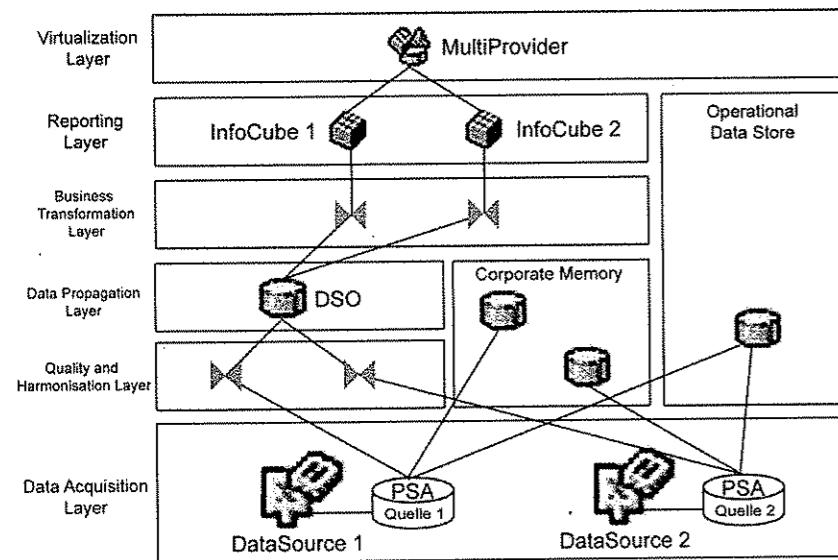


Abb. 3 - LSA nach SAP (vgl. Wolf & Yamada 2010, S.166)

Die hellgrau hinterlegten Schichten in Abb. 3 stellen das Enterprise Data Warehouse (EDW) dar. Es ist unabhängig davon, welche Anwendung zur Analyse der Daten zugrunde liegt; diese Schichten verschaffen einen allgemeingültigen Blick auf die Daten. Die hellen Schichten hingegen sind von der Anwendung abhängig und dementsprechend verschieden. Sie werden nach den Anforderungen modelliert und erstellt.

Die nachfolgende Abb. 4 zeigt die bekannte LSA, dieses Mal mit den entsprechenden Objekten des SAP BW, mit denen die Modellierung umgesetzt wird.

Der Data Acquisition Layer wird mit Hilfe der Data Source umgesetzt. Die Rohdaten werden hier in der PSA (Persistent Staging Area) gespeichert bis sie weiterverarbeitet wurden. Dabei handelt es sich um eine Tabelle mit der Struktur der DataSource. Der Prozess im Quality and Harmonization- sowie Business Transformation Layer erfolgt durch Transformationen. In diesen wird festgelegt, welche Felder aus der vorigen Schicht in die nächste übernommen werden soll und ob dabei etwas berechnet oder verändert werden muss. Der Data Propagation Layer wird durch das DataStore-Objekt umgesetzt. Den Anforderungen der Optimierung für Analysen entspricht der InfoCube, weshalb dieser für den Reporting Layer genutzt wird. Das Corporate Memory wird, wie der Data Propagation Layer, mit einem DSO umgesetzt. Hier handelt es sich aber um eine besondere Form, um das schreiboptimierte DSO. Es wird deshalb eingesetzt, da die Daten aus der PSA so schnell wie möglich weggeschrieben werden sollen, in diesem Fall in das Corporate Memory. Das wiederholte Laden aus einem schreiboptimierten DSO ist nicht vorgesehen, weshalb es auch nur selten durchgeführt werden sollte. Beim Operational Data Store werden die Daten wie beim Corporate Memory nicht harmonisiert oder transformiert. Sie stehen in ihrer Rohform für Abfragen zur Verfügung. Sind zum Zeitpunkt der Abfrage nicht alle aktuellen Daten geladen, wird dies direkt über die DataSource getan. Diese Art der Abfragen ist nur bei wichtigen Echtzeitdaten zu empfehlen, da dabei auch jedes Mal eine Datenabfrage im operativen Quellsystem angestoßen wird. Die grünen, zueinander gekehrten Dreiecke stellen eine Transformation dar.



BW Objekten (Eigene Abbildung)

Abb. 4 - LSA mit SAP