



Helmut Tack

Performance-Optimierung für SAP® BW

- ▶ Performancemessungen und deren Interpretation
- ▶ Archivierung und regelmäßige Bereinigung der Datenbank
- ▶ 30 praktische Tipps zur Optimierung auf Basis SAP BW 7.3
- ▶ Praxisnahes Nachschlagewerk für den täglichen Betrieb

Inhaltsverzeichnis

Einleitung	9
1 Grundlegende Betrachtungen	13
2 Grundlagen der SAP BW-Datenhaltung	17
2.1 Flache Datenspeicher	21
2.2 Mehrdimensionale Datenspeicher	37
2.3 Virtuelle und temporäre Datenlieferanten	44
2.4 Schlüsselbildung innerhalb der relationalen Datenbank und in klassischen Sternschema	58
2.5 Schlüsselbildung innerhalb SAP BW-Datenobjekten	61
3 Datenmodellierung aus Sicht der Performance	75
3.1 Die Extraktionsmodellierung	75
3.2 Die InfoCube-Modellierung	91
3.3 Die DSO-Modellierung	106
3.4 Spezielle Datenmodelle	117
3.5 Die Query-Modellierung	121
4 Performancemessungen und deren Interpretation	127
4.1 Systeminformationen	128
4.2 Datenbank-Messungen	130
4.3 Hardware-Messungen	136
4.4 Prozessübergreifende Messungen	143
4.5 Performancemessungen für BW-Objekte	152
4.6 Performancemessungen für Reportingobjekte	156
4.7 Schnelle Störungssuche im BW-System	164

5 Performanceoptimierung	167
5.1 Reorganisieren von BEx Web Application Bookmarks	167
5.2 Reorganisieren von BEx Broadcaster Bookmarks	171
5.3 Reorganisieren von Rollen und Favoriten-Arbeitsmappen	174
5.4 Aggregate-Checks durchführen	176
5.5 Anwendungsprotokolle löschen	192
5.6 Archivierung und Reorganisation von IDocs	194
5.7 Verknüpfungen mit IDocs löschen	199
5.8 Löschen von alten Meldungen, Parametern und temporären DTP-Daten	202
5.9 Archivierung von Request-Verwaltungsdaten	204
5.10 Archivierung oder Löschen von BI-Berechtigungsprotokollen	207
5.11 Archivierung von Änderungssätzen der Analyseberechtigungen	212
5.12 Löschen von Fehlerprotokollen ohne PSA	213
5.13 Prüfen und Bereinigen von PSA-Tabellen	214
5.14 DTP-Fehlersätze löschen	218
5.15 Löschen alter Prozessketten-Protokolle	220
5.16 Löschen alter Job-Protokolle	226
5.17 Löschen alter Job-Protokolle ohne Varianten	232
5.18 Reorganisation dunkler Mappen	233
5.19 Prüfen und Bereinigen von Tabellenänderungsprotokollen	240
5.20 Prüfen und Löschen alter TRFC-Queues	242
5.21 Löschen temporärer Datenbankobjekte	245
5.22 Löschen leerer F-Partitionen	247
5.23 Reorganisation der TemSe und des Spools	248
5.24 InfoCube-Komprimierung	251

5.25	Löschen überflüssiger BW-Statistiktabellen	253
5.26	Löschen verwaister Job-Logs	255
5.27	Löschen von RSTT-Traces	257
5.28	Löschen unbenutzter Stammdaten	258
5.29	Reparieren sekundärer PSA-Indizes	262
5.30	Reorganisation und Reparatur fehlender und fehlerhafter Indizes	264
5.31	Bereinigung des Filesystems	266
6	Möglichkeiten der Archivierung	271
7	Housekeeping-Konzept	275
7.1	Tägliche Maßnahmen	275
7.2	Wöchentliche Maßnahmen	277
7.3	Monatliche Maßnahmen	280
7.4	Jährliche Maßnahmen	281
8	Zusammenfassung und Ausblick	285
9	Glossar	287
A	Über den Autor	303
B	Index	305
C	Disclaimer	309
	Weitere Bücher von Espresso Tutorials	310

2 Grundlagen der SAP BW-Datenhaltung

Um SAP BW-Systeme optimieren zu können, ist ein Verständnis der Datenhaltung unerlässlich. In diesem Kapitel werde ich Ihnen die häufigsten Arten der Datenhaltung und die entscheidenden Unterschiede zu anderen datenbankbasierten Systemen, insbesondere zum klassischen Data Warehouse, erläutern.

Innerhalb einer *EDV* werden – wie der Name schon sagt – Daten auf elektronischem Wege verarbeitet. Wenn es sich dabei um Daten handelt, die nach der Verarbeitung noch anderen Vorgängen oder auch demselben Prozess zur Verfügung stehen sollen, ist eine Datenbank erforderlich.

In der Regel folgen heutige Datenbanken dem *Relationsprinzip nach Codd*. (Edgar Frank Codd, Entwickler des Relational Database Management Systems RDBMS).

Nach diesem Prinzip besteht zwischen redundant verwendbaren Objekten und Informationen eine Primär-Fremdschlüssel-Beziehung, die dazu geeignet ist, Datenbanken schlanker, eindeutiger und wiederverwendbarer zu machen.

Datenbanken werden unter Verwendung der sogenannten *Normalisierungsregeln* aufgebaut. Sie bieten einen eindeutigen Schlüssel auf einen oder mehrere der zum Schlüssel gehörenden Werte.

Innerhalb des SAP BW-Systems kommen sowohl normalisierte als auch nicht normalisierte Datenbankmodelle zur Anwendung.

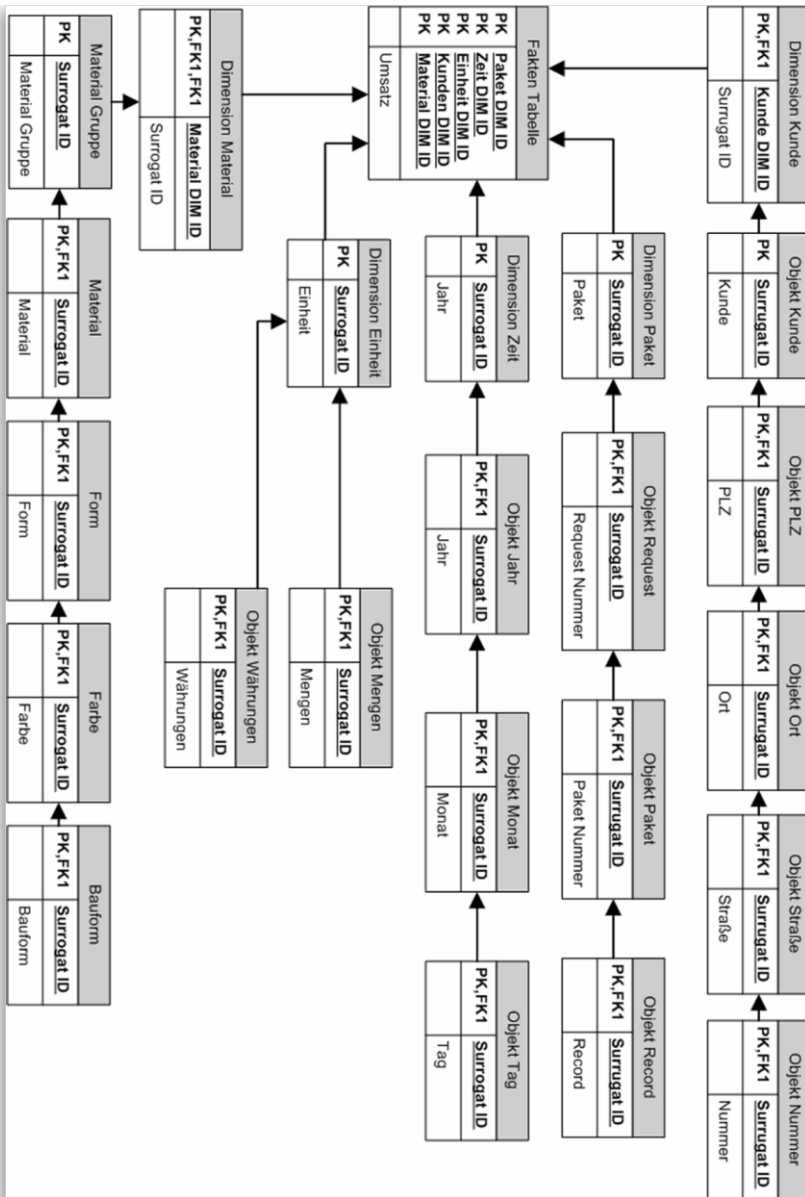


Abbildung 2.1: Denormalisierte Dimensionstabellen verweisen auf die Faktentabelle

Normalisierungskonzept des InfoCubes



Während die *Dimensionstabellen* eines InfoCubes *denormalisiert* sind und sein müssen, sind dessen *Faktentabellen hoch normalisiert* (siehe Abbildung 2.1). Nur so ist es möglich, dass jedwede Kombination von Merkmalen auf einen oder mehrere Kennzahlwert(e) verweisen kann.

SAP BW unterscheidet grundsätzlich zwischen

- ▶ *persistenter Datenhaltung* (Datenziele) = physisch auf der Datenbank und
- ▶ *transienter Datenbereitstellung* (Nicht-Datenziele) = zur Laufzeit verfügbare Daten.

Eine weitere wesentliche Unterscheidung innerhalb der Datenhaltung des SAP BW besteht in der flachen und der mehrdimensionalen Datenhaltung:

- ▶ Merkmale, DataStore-Objekte (DSO) = *flache Datenhaltung*
- ▶ InfoCubes = *mehrdimensionale Datenhaltung*.

Lassen Sie uns in Abbildung 2.2 etwas genauer darauf eingehen. Zunächst zur Erinnerung:

Datenziele:

- ▶ sind InfoProvider,
- ▶ belegen physische Speicher auf der Datenbank,
- ▶ können zum Reporting herangezogen werden.

Nicht-Datenziele:

- ▶ sind InfoProvider,
- ▶ belegen keinen physischen Speicher auf der Datenbank,
- ▶ können zum Reporting herangezogen werden.

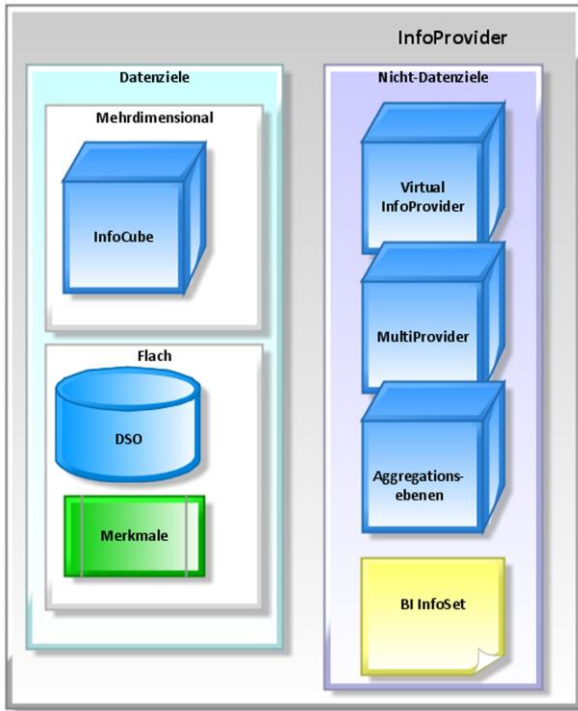


Abbildung 2.2: Infoprovidertypen

Datenmodell und Performance



Die Art der Datenhaltung und Datenbereitstellung hat erheblichen Einfluss auf die Performance von Berichten. Wählen Sie bei der Modellierung nicht das Modell aus, welches nur die Datenbank schont, sondern das für den auszuführenden Bericht optimale.

Mit der neuesten Technologie aus dem Hause SAP findet aus Sicht der Datenhaltung ein Quantensprung statt: SAP HANA stellt persistente Daten durch einen transienten Prozess im aktiven Speicher als Datenbank zur Verfügung, was die Bereitstellungszeiten erheblich verringert.

2.1 Flache Datenspeicher

Bei flachen Datenspeichern handelt es sich um BW-Objekte, die sich technisch als einfache Tabellen mit Kopf und Zeilen darstellen lassen. Im Kopf befinden sich die Felder, über die Tabellenspalten gebildet werden, während sich in den Zeilen die gespeicherten Informationen befinden.

Diese Tabellen können mit anderen in einer Relation stehen (siehe Abbildung 2.4). Besteht eine Beziehung zu einer oder mehreren anderen Tabellen auf Schlüsselebene, sprechen wir von einer *Primär-Fremdschlüsselbeziehung*. Die Fremdschlüssel beschaffen (verweisen auf) Werte in der verbundenen Tabelle.

Manche Tabellen stehen allerdings nicht in Relation zueinander, sondern bilden mit anderen Tabellen einen speziellen Modelltyp. Wie sich diese Beziehungen aufbauen können, soll Ihnen die folgende einfache Grafik verdeutlichen (siehe Abbildung 2.3).

Aus einer Reihe von Datenbanktabellen werden Informationen verwendet, um einen flüchtigen Datenwert (Bewegungsdatum) innerhalb eines Datenspeichers zu adressieren. Adressieren ist meiner Ansicht nach in diesem Zusammenhang der wohl beste Ausdruck, da die Kombination der verwendeten Stammdatenschlüssel eine eindeutige Datenbankadresse bildet. Dieser Schlüssel ist wiederum normalisiert und eindeutig.

Tabelle	KundePK	OrtPK	StraßePK	NummerPK	MaterialPK	Umsatz
ZCUST	1022	24	66	124	00010223824	400,00 €
ZCUST	1031	24	66	124	00010223331	1.254,23 €

Abbildung 2.3: Adressierung über Schlüssel

Wie Sie in der Abbildung sehen, befinden sich aus Sicht der Datenbank an derselben Adresse zwei verschiedene Kunden, die unterschiedliche Umsätze generiert haben.

Über ein solches Konstrukt werden in der Regel Stammdaten am InfoProvider miteinander kombiniert und somit die Bewegungsdaten eindeutig den Stammdaten zugeordnet. Auf Stammdatenebene las-

B Index

A

- Activation Queue 66, 67, 69
 - Definition 32
 - eindeutige Sätze 112
- Active Data 32, 67, 69, 70
- Aggregat 37, 43, 98, 99, 158, 176
 - Aggregationsselektion 177
 - Bewertung 187
 - Check 176
 - Definition 287
 - Indexreparatur 264
 - Letzte Nutzung 188
 - Letzter Rollup 188
 - Messungen 161
 - Prüfung 181, 186, 189
 - Störungen 188
 - Verdichtung 188
 - Vorschläge 253
- Aggregationsverfahren 32, 49
- Aktivierungsmenge 104
- Application Program Interface 70
- Ausprägung 63, 96, 117, 122, 177, 183, 260

B

- Bewegungsdaten 13, 31, 37, 72, 131
- BEx Query Designer 123

- BI-InfoSet 45
 - Definition 53
- Bookmark 169
 - Definition 167
 - verwalten 172

C

- Cache-Monitor 159
- CATT 258
- Change Log 32, 33, 34, 35, 36, 66, 67, 69, 112
- Change Run *Siehe* Stammdaten
- Computer Aided Test Tool *Siehe* CATT

D

- Data Dictionary 134, 215
 - Qualität 135
- Data Mining
 - DSO im Data Mining 70
- Data Warehouse 13, 253
- DataSource 80, 84, 85, 89
- DataSources 194
- DataStore Objekt 108, 114
- DataStore Objekte 19, 31, 106
- Dateisystem 142, 248
- Datenbankindex

Definition 290

Datenrequests *Siehe* Request

Datenziel 19

DDIC *Siehe* Data Dictionary

degenerierte Dimension
 Siehe Line Item Dimension

denormalisiert 19

Dimensions ID
 Definition 290

Dimensionstabellen 19, 60,
 61, 251
 Definition 290
 Namenskonvention 251

DIM-ID 28, 62, 96

DSO 45, 58, 106, 215, 276,
 Siehe DataStore Objekt

direktes Schreiben 31, 70

Indizierung 109

Laufzeitparameter 114

schreiboptimiert 31, 69

Standard 31

transaktionales 70

DTP *Siehe* Daten Transfer

 Prozess

dunkle Mappe 233
 Definition 291

E

Extraktion 16

Extraktor 80, 168, 298

F

Faktentabelle 19, 39, 41, 60,
 61, 73, 93, 118, 120, 247
 Definition 291

Namenskonvention 251

Filesystem *Siehe* Dateisystem

flache Datenhaltung 19

Funktionsbaustein 76, 163,
 243

Funktionsgruppe 163

G

GUI
 Zeit 150

H

Hierarchie- und
 Attributsänderungslauf
 Siehe Change Run

hoch normalisiert 19

Housekeeping 255, 275

I

IDoc 80, 86, 89, 91, 194, 197,
 199
 Definition 292

Impact Analyse 179, 183
 Definition 292

Index *Siehe* Datenbankindex

 Clustering 90

 Datenbankindex 125

 Neuaufbau 264

 Partitionierung 215

 Primärindex 109

 Probleme 214

 Reparatur 277, 280

 Sekundärindex 109, 262

InfoObjekt 30, 177, 178, 207,
 258, 262

Definition 292
 InfoPackage 90
 Definition 292
 InfoProvider 13, 15, 16, 19,
 21, 28, 44, 45, 55, 71, 72,
 90, 91, 106, 122, 123, 160,
 171, 186, 207, 214, 253,
 259, 273
 Definition 292
 Information Broadcasting
 E-Mail Bursting 171

K

Kennzahl 15
 Konsistenz
 vertikale und horizontale 66

L

Line Item 62, 93, 97
 Definition 293

M

mehrdimensionale
 Datenhaltung 19
 Metadaten 131, 167, 168
 Definition 293
 Metakette 148, 221, 279
 Definition 293
 MetaKette 281
 Microsoft Excel 13, 175
 Mini-InfoCube *Siehe* Aggregat
 MultiCube *Siehe*
 MultiProvider
 MultiProvider 45, 156, 298,
 299

Definition 293

N

Nearline Storage 259, 272,
 282
 Definition 294
 NLS *Siehe* Nearline Storage
 Normalisierungsregeln 17

O

OLAP
 Cache 122, 159, 161
 Laufzeit 125
 Prozessor 46, 188
 Prozesszeit 55
 OLAP-Prozessor
 Definition 294

P

persistent 19, 44, 45, 55
 Definition 294
 Persistent Staging Area 69,
 89
 Definition 294
 Primär-Fremdschlüssel-
 Beziehung 17
 Definition 294
 PSA *Siehe* Persistent Staging
 Area

Q

Querymonitor 49
 Definition 295
 Querystatistik 50

Definition 295

R

Relationsprinzip nach Codd
17

Remote Function Call 83, 242
Definition 295

Reparaturrequest *Siehe*
Request

Request 23, 32, 41, 68, 75,
80, 81, 86, 87, 204, 213,
214, 218, 244, 247, 271,
276, 280
Delta 83

request-partitioniert *Siehe*
Request

RFC *Siehe* Remote Function Call

Rolle 174
Definition 296
IDoc 199

S

SAP HANA 11, 20, 153

Schlüsselbeziehung 21

Service API 76, 78
Definition 297

Stammdaten 13, 23, 24, 60,
61, 117

Attribute 25

Ausprägungen 26

Change Run 177

Datenbank 130

doppelte Sätze 219

Schlüssel 21, 25

SID-Ermittlung 106

Surrogat-Tabelle 25

Tabelle 25

Trägermerkmal 23, 25

T

temporaler Join 53, 54

Definition 298

Transferstruktur 80, 168, 194

Definition *Siehe*

Transformation 31, 56, 88

Definition 299

Routinen 219

transient 19

V

virtueller Speicher 141

Z

Zombie-Jobs

Bereinigung 250

Definition 250