

Inhaltsverzeichnis

1 Business-Intelligence-Architektur	1
1.1 Data Warehouse	1
1.2 OLAP und mehrdimensionale Datenbanken	4
1.3 Architekturvarianten	6
1.3.1 Stove-Pipe-Ansatz	6
1.3.2 Data Marts mit abgestimmten Datenmodellen	7
1.3.3 Core Data Warehouse	8
1.3.4 Hub-and-Spoke-Architektur	10
1.3.5 Data-Mart-Busarchitektur nach Kimball	12
1.3.6 Corporate Information Factory nach Inmon	13
1.3.7 Architekturvergleich Kimball und Inmon	15
1.4 Schichtenmodell der BI-Architektur	16
1.4.1 Acquisition Layer	18
1.4.2 Integration Layer	20
1.4.3 Reporting Layer	21
1.4.4 Modellierung im Schichtenmodell	22
2 Mehrdimensionale Datenstrukturen	25
2.1 Datenmodelle und Datenmodellierung	25
2.2 Grundbestandteile mehrdimensionaler Datenstrukturen	28
2.3 Hierarchische Dimensionsstrukturen	33
2.3.1 Strukturlose Dimensionen	35
2.3.2 Balancierte Baumstrukturen	35
2.3.3 Balancierte Waldstrukturen	36
2.3.4 Unbalancierte Baum- und Waldstrukturen	37
2.3.5 Parallelle Hierarchien	37
2.3.6 Heterarchien (Many-Many-Beziehungen)	38

2.3.7	Rekursive Hierarchien und bebuchbare Knoten	39
2.3.8	Hierarchieattribute	40
2.4	Kennzahlen und deren Berechnung	43
2.4.1	Kennzahlen und Kennzahlensysteme	43
2.4.2	Kennzahlen im mehrdimensionalen Modell	47
2.4.3	Additivitätseigenschaft	49
2.5	Historisierung und Zeitabhängigkeit	49
3	Semantische mehrdimensionale Modellierung	53
3.1	Methoden auf Basis der Entity-Relationship-Modellierung	53
3.1.1	Grundbestandteile der ER-Modellierung	54
3.1.2	Erweiterte ERM-Konstrukte	57
3.1.3	ER-basierte mehrdimensionale Modellierung	61
3.1.4	Mehrdimensionales ER-Modell (ME/R)	62
3.2	Mehrdimensionale Modellierung mit ADAPT	64
3.2.1	Dimensionsmodellierung in ADAPT	64
3.2.2	Varianten der Hierarchiemodellierung	81
3.2.3	Modellierung von Würfeln	85
3.3	T-ADAPT: Modellierung von Zeitabhängigkeit	88
4	Bestandteile und Varianten des Star-Schemas	93
4.1	Einfaches Star-Schema	94
4.1.1	Grundform des Star-Schemas	94
4.1.2	Abbildung von Kennzahlen und Kennzahlensystemen	99
4.1.3	Attribute in Dimensionen	101
4.2	Modellierung von Dimensionshierarchien	102
4.2.1	Flache Strukturen	102
4.2.2	Balancierte Baum- und Waldstrukturen	102
4.2.3	Unbalancierte Strukturen	103
4.2.4	Parallele Hierarchien	104
4.2.5	Anteilige Verrechnung und Heterarchien	104
4.3	Normalisierung von Dimensionen	105
4.4	Übergang von T-ADAPT zum logischen Modell	107
4.4.1	Transformation von Dimensionen	107
4.4.2	Abbildung von Attributen	109
4.4.3	Transformation von Scopes	110
4.4.4	Behandlung spezieller ADAPT-Varianten	114

4.5	Modellierung von Parent-Child-Hierarchien	116
4.5.1	Iterative Abfrage	118
4.5.2	Einstufige Rekursion	119
4.5.3	Mehrstufige Rekursion	120
4.5.4	Rekursives SQL	121
4.5.5	Brückentabellen	124
5	Historisierung und Zeitabhängigkeit im Data Warehouse	129
5.1	Historisierung im Star-Schema	130
5.1.1	Keine Historisierung bei Type 0 und Type 1	132
5.1.2	Type-3-Attribut-Paare	133
5.1.3	Versionen und Zeitstempelung für as is und as of	134
5.2	Bewegungsdatensicht in der Historisierung	138
5.2.1	As-posted-Type-2-Szenario	138
5.2.2	Snapshot-Verfahren	142
5.2.3	Vollständige Zeitstempelung plus as posted	145
5.2.4	Varianten für hybride Historisierung	147
5.3	Best Practices der Historisierung	150
5.4	Bitemporale Historisierung	151
6	Dimensionsmodellierung	153
6.1	Dimensionstabellen	153
6.1.1	Degenerierte Dimensionen	153
6.1.2	Housekeeping und technische Dimensionen	155
6.1.3	Große Dimensionen	156
6.1.4	Mehrsprachigkeit	158
6.1.5	Outrigger-Tabellen	159
6.2	Rollen von Dimensionen	163
6.3	Many-Many-Beziehungen	166
6.3.1	Hierarchien über Faktentabellen	167
6.3.2	Mehrwertige Dimensionen (multi valued dimensions) ...	170
6.3.3	Many-Many-Beziehungen über Dimensionen	171
6.3.4	Mehrwertige Attribute	173
6.4	Datum- und Zeitdimension	174

7 Faktenmodellierung	181
7.1 Kennzahlen und Kennzahlensysteme	181
7.2 Aggregate	188
7.3 Snowflake-Schema	191
7.4 Faktenlose Faktentabellen	194
7.5 Granularität	196
7.6 Additivität und berechnete Kennzahlen	200
7.6.1 Transaktionsfaktentabellen	200
7.6.2 Bestandsmodelle	202
7.6.3 Prozessmodelle	207
7.7 Abgeleitete Schemata	209
8 Core-Data-Warehouse-Modellierung	213
8.1 Aufgaben der Data-Warehouse-Komponenten	214
8.1.1 Datenintegrations-Framework	214
8.1.2 Aufgaben und Komponenten in Multi-Layer-Architekturen	216
8.1.3 Eignungskriterien für Methoden der Core-Data-Warehouse-Modellierung	219
8.2 Star-Schema-Modellierung im Core Data Warehouse	221
8.2.1 Granulare Star-Schemata im Core Data Warehouse	221
8.2.2 Bewertung dimensionaler Core-Data-Warehouse-Modelle	223
8.3 3NF-Modelle im Core Data Warehouse	224
8.3.1 Core-Data-Warehouse-Modellierung in 3NF	224
8.3.2 Historisierungsaspekte von 3NF-Modellen	225
8.3.3 Bewertung der 3NF-Modellierung im Core Data Warehouse	226
8.4 Data-Vault-Ansatz	227
8.4.1 Hub-Tabellen	228
8.4.2 Satellite-Tabellen	229
8.4.3 Link-Tabellen	234
8.4.4 Zeitstempel im Data Vault	237
8.4.5 Harmonisierung von fachlichen Schlüsseln	238
8.4.6 Agilität in Data-Vault-Modellen	239
8.4.7 Vorgehensweise zur Data-Vault-Gestaltung	241
8.4.8 Bewertung der Data-Vault-Methode	242

Anhang	245
A Abkürzungen	247
B Literaturverzeichnis	249
Index	255