

# Inhaltsverzeichnis

<b>1 Einführung</b>	1
1.1 Ziele der Vorlesung	1
1.2 Die Idee	1
1.3 Lernkarte	2
1.4 Selbsttest-Frage	3
Literaturhinweise	3
<b>Teil I Die Zukunft von Enterprise-Computing</b>	5
<b>2 Neue Anforderungen an Enterprise Computing</b>	7
2.1 Verarbeitung von Ereignisdaten	7
2.1.1 Sensordaten	7
2.1.2 Analyse von Spiel-Ereignissen	8
2.2 Kombination von strukturierten und unstrukturierten Daten	9
2.2.1 Patientendaten	10
2.2.2 Flugzeugwartungs-Berichte	10
2.3 Soziale Netzwerke und das Web	11
2.4 Cloud-Anwendungen	11
2.5 Mobile Anwendungen	12
2.6 Produktions- und Vertriebsplanung	13
2.6.1 Produktionsplanung	13
2.6.2 Der Available-to-Promise-Check	13
2.7 Selbsttest-Frage	13
Literaturhinweise	14
<b>3 Merkmale von Unternehmensanwendungen</b>	15
3.1 Vielfältige Anwendungen	15
3.2 OLTP versus OLAP	15
3.3 Nachteile der Trennung von OLAP und OLTP	16
3.4 Mythos OLTP- versus OLAP-Zugriffsmuster	16
3.5 Kombinieren von OLTP- und OLAP-Daten	17
3.6 Merkmale von Unternehmensdaten	17
3.7 Selbsttest-Frage	18
Literaturhinweise	18

<b>4 Wandel in der Hardware</b> . . . . .	19
4.1 Speicherzellen . . . . .	19
4.2 Speicherhierarchie . . . . .	20
4.3 Cache im Detail . . . . .	21
4.4 Addressübersetzung . . . . .	22
4.5 Prefetching . . . . .	23
4.6 Speicherhierarchie und Latenzzeiten . . . . .	23
4.7 Non-Uniform Memory-Architektur . . . . .	25
4.8 Skalieren von Hauptspeicher-Systemen . . . . .	26
4.9 Remote Direct Memory Access (RDMA) . . . . .	28
4.10 Selbsttest-Fragen . . . . .	28
Literaturhinweise . . . . .	28
<b>5 SanssouciDB – Ein Entwurf für eine In-Memory Datenbank</b> . . . . .	31
5.1 Datenspeicherung im Hauptspeicher . . . . .	31
5.2 Spaltenorientierung . . . . .	31
5.3 Auswirkungen der Spaltenorientierung . . . . .	32
5.4 Aktive und Passive Daten . . . . .	33
5.5 Überblick über die Architektur . . . . .	33
5.6 Selbsttest-Fragen . . . . .	33
Literaturhinweis . . . . .	34
<b>Teil II Grundlagen der Datenbankspeichertechniken</b> . . . . .	35
<b>6 Wörterbuch-Codierung</b> . . . . .	37
6.1 Kompressionsbeispiel . . . . .	38
6.1.1 Beispiel für Wörterbuch-Codierung: Vornamen . . . . .	38
6.1.2 Beispiel für Wörterbuch-Codierung: Geschlecht . . . . .	39
6.2 Sortierte Wörterbücher . . . . .	40
6.3 Operationen mit codierten Werten . . . . .	40
6.4 Selbsttest-Fragen . . . . .	41
<b>7 Kompression</b> . . . . .	43
7.1 Präfix-Encoding . . . . .	43
7.1.1 Beispiel . . . . .	44
7.2 Run-Length Encoding . . . . .	45
7.2.1 Beispiel . . . . .	46
7.3 Cluster-Encoding . . . . .	46
7.3.1 Beispiel . . . . .	47
7.4 Indirect-Encoding . . . . .	48
7.4.1 Beispiel . . . . .	49
7.5 Delta-Encoding . . . . .	51
7.5.1 Beispiel . . . . .	51
7.6 Einschränkungen . . . . .	52
7.7 Selbsttest-Fragen . . . . .	53
Literaturhinweis . . . . .	54

<b>8 Datenlayout im Hauptspeicher</b> . . . . .	55
8.1 Effekte von Caching auf die Leistung von Anwendungen . . . . .	55
8.1.1 Das Schrittweiten-Experiment . . . . .	55
8.1.2 Das Größen-Experiment . . . . .	57
8.2 Zeilen- und spaltenorientierte/-basierte Layouts . . . . .	59
8.3 Vorteile eines Spalten-Layouts . . . . .	61
8.4 Hybride Tabellen-Layouts . . . . .	62
8.5 Selbsttest-Frage . . . . .	62
Literaturhinweise . . . . .	63
<b>9 Partitionierung</b> . . . . .	65
9.1 Definition und Klassifikation . . . . .	65
9.2 Vertikale Partitionierung . . . . .	65
9.3 Horizontale Partitionierung . . . . .	66
9.4 Wahl einer geeigneten Partitionierungsstrategie . . . . .	68
9.5 Selbsttest-Fragen . . . . .	68
Literaturhinweis . . . . .	69
<b>Teil III In-Memory Datenbank-Operatoren</b> . . . . .	71
<b>10 Löschen von Daten: DELETE</b> . . . . .	73
10.1 Beispiel für physisches Löschen . . . . .	73
10.2 Selbsttest-Fragen . . . . .	75
Literaturhinweis . . . . .	75
<b>11 Einfügen von Daten: INSERT</b> . . . . .	77
11.1 Beispiel: Einfügen eines neuen Tupels . . . . .	77
11.1.1 Einfügen ohne neuen Wörterbuch-Eintrag . . . . .	77
11.1.2 Einfügen mit neuem Wörterbuch-Eintrag . . . . .	78
11.2 Überlegungen zur Leistung . . . . .	81
11.3 Selbsttest-Fragen . . . . .	82
<b>12 Aktualisieren von Einträgen – UPDATE</b> . . . . .	85
12.1 Arten von Updates . . . . .	85
12.1.1 Aggregat-Updates . . . . .	85
12.1.2 Status-Updates . . . . .	85
12.1.3 Wert-Updates . . . . .	86
12.2 Update-Beispiel . . . . .	86
12.3 Selbsttest-Fragen . . . . .	88
Literaturhinweise . . . . .	89
<b>13 Tupel-Rekonstruktion</b> . . . . .	91
13.1 Einführung . . . . .	91
13.2 Tupel-Rekonstruktion in zeilenorientierten Datenbanken . . . . .	91
13.3 Tupel-Rekonstruktion in spaltenorientierten Datenbanken . . . . .	92
13.4 Weitere Beispiele und Diskussion . . . . .	93
13.5 Selbsttest-Fragen . . . . .	94

<b>14 Scan-Leistung</b> . . . . .	97
14.1 Einführung . . . . .	97
14.2 Zeilen-Layout: vollständiger Tabellen-Scan . . . . .	98
14.3 Zeilen-Layout: Schrittweiten-Zugriff . . . . .	98
14.4 Spaltenorientiertes Layout: vollständiger Spalten-Scan . . . . .	99
14.5 Weitere Beispiele und Diskussion . . . . .	100
14.6 Selbsttest-Frage . . . . .	101
<b>15 Abfragen von Einträgen – SELECT</b> . . . . .	103
15.1 Relationale Algebra . . . . .	103
15.1.1 Kartesisches Produkt . . . . .	103
15.1.2 Projektion . . . . .	103
15.1.3 Selektion . . . . .	104
15.2 Abfrageausführung . . . . .	104
15.3 Selbsttest-Fragen . . . . .	106
<b>16 Materialisierungsstrategien</b> . . . . .	109
16.1 Aspekte der Materialisierung . . . . .	109
16.2 Beispiel . . . . .	110
16.3 Frühe Materialisierung . . . . .	111
16.4 Späte Materialisierung . . . . .	113
16.5 Selbsttest-Fragen . . . . .	116
Literaturhinweise . . . . .	116
<b>17 Parallele Datenverarbeitung</b> . . . . .	117
17.1 Hardware-Schicht . . . . .	117
17.1.1 Multi-Core-CPUs . . . . .	118
17.1.2 Single Instruction Multiple Data (SIMD) . . . . .	119
17.2 Software-Schicht . . . . .	121
17.2.1 Amdahls Gesetz . . . . .	121
17.2.2 Shared Memory . . . . .	122
17.2.3 Message Passing . . . . .	122
17.2.4 MapReduce . . . . .	123
17.3 Selbsttest-Fragen . . . . .	123
Literaturhinweise . . . . .	124
<b>18 Indizes</b> . . . . .	125
18.1 Indizes: Ein Ansatz zur Optimierung von Abfragen . . . . .	125
18.2 Technische Betrachtungen . . . . .	125
18.3 Invertierter Index . . . . .	127
18.4 Diskussion . . . . .	130
18.4.1 Speicherverbrauch . . . . .	130
18.4.2 Abfrage-Leistung . . . . .	131
18.5 Selbsttest-Fragen . . . . .	133
Literaturhinweis . . . . .	133

<b>19 JOIN</b> . . . . .	135
19.1 Ausführung von Joins im Hauptspeicher . . . . .	136
19.2 Hash-Join . . . . .	137
19.2.1 Beispiel: Hash-Join . . . . .	138
19.3 Sort-Merge-Join . . . . .	139
19.3.1 Beispiel: Sort-Merge-Join . . . . .	140
19.4 Die Wahl eines Join-Algorithmus . . . . .	141
19.5 Selbsttest-Fragen . . . . .	142
<b>20 Aggregatfunktionen</b> . . . . .	145
20.1 Aggregation am Beispiel der Funktion COUNT . . . . .	145
20.2 Selbsttest-Fragen . . . . .	147
<b>21 Paralleles SELECT</b> . . . . .	149
21.1 Parallelisierung . . . . .	149
21.2 Selbsttest-Frage . . . . .	152
<b>22 Workload-Management und Scheduling</b> . . . . .	153
22.1 Die Macht der Geschwindigkeit . . . . .	153
22.2 Scheduling . . . . .	154
22.3 Management gemischter Workloads . . . . .	154
22.4 Selbsttest-Fragen . . . . .	155
Literaturhinweis . . . . .	155
<b>23 Paralleler Join</b> . . . . .	157
23.1 Teilweise parallelisierter Hash-Join . . . . .	157
23.2 Paralleler Hash-Join . . . . .	158
23.3 Selbsttest-Frage . . . . .	159
Literaturhinweise . . . . .	159
<b>24 Parallele Aggregation</b> . . . . .	161
24.1 Aggregatfunktionen . . . . .	161
24.2 Parallel Aggregation unter Verwendung einer Hash-Funktion . . . . .	161
24.3 Selbsttest-Fragen . . . . .	164
Literaturhinweis . . . . .	164
<b>Teil IV Fortgeschrittene Datenbank-Speichertechniken</b> . . . . .	165
<b>25 Differential Buffer</b> . . . . .	167
25.1 Das Konzept . . . . .	167
25.2 Die Implementierung . . . . .	167
25.3 Lebensdauer von Tupeln . . . . .	169
25.4 Selbsttest-Fragen . . . . .	170
Literaturhinweise . . . . .	171

<b>26 Insert-Only</b> . . . . .	173
26.1 Definition des Insert-Only-Ansatzes . . . . .	173
26.2 Punkt-Repräsentation . . . . .	174
26.3 Intervall-Repräsentation . . . . .	176
26.4 Concurrency Control: Snapshot-Isolation . . . . .	177
26.5 Insert-Only: Vorteile und Herausforderungen . . . . .	179
26.6 Selbsttest-Fragen . . . . .	179
Literaturhinweis . . . . .	180
<b>27 Der Merge-Prozess</b> . . . . .	181
27.1 Der asynchrone Online-Merge-Prozess . . . . .	182
27.1.1 Merge-Vorbereitungsphase . . . . .	183
27.1.2 Attribut-Merge-Phase . . . . .	183
27.1.3 Commit-Merge-Phase . . . . .	184
27.2 Beispielhafter Attribut-Merge einer Spalte . . . . .	184
27.3 Optimierungen der Merge-Operation . . . . .	187
27.3.1 Verwendung des Wörterbuches des Main Stores . . . . .	187
27.3.2 Single Column Merge . . . . .	187
27.3.3 Vereinheitlichtes Tabellen-Konzept . . . . .	188
27.4 Selbsttest-Fragen . . . . .	189
Literaturhinweise . . . . .	189
<b>28 Logging</b> . . . . .	191
28.1 Logging-Infrastruktur . . . . .	191
28.2 Logisches versus Wörterbuch-codiertes Logging . . . . .	193
28.3 Beispiel . . . . .	196
28.4 Selbsttest-Fragen . . . . .	198
Literaturhinweise . . . . .	199
<b>29 Wiederherstellung</b> . . . . .	201
29.1 Lesen der Metadaten . . . . .	201
29.2 Wiederherstellen der Datenbank . . . . .	202
29.3 Selbsttest-Fragen . . . . .	202
Literaturhinweis . . . . .	203
<b>30 On-the-Fly-Datenbankreorganisation</b> . . . . .	205
30.1 Reorganisation in einem Row Store . . . . .	205
30.2 On-the-Fly-Reorganisation in einem Column Store . . . . .	206
30.3 Exkurs: Multi-Tenancy erfordert Online-Reorganisation . . . . .	206
30.4 Aktive und passive Daten . . . . .	208
30.5 Selbsttest-Fragen . . . . .	209
Literaturhinweise . . . . .	211

<b>Teil V Grundlagen für die Entwicklung neuer Enterprise Anwendungen . . . . .</b>	213
<b>31 Auswirkungen auf die Anwendungsentwicklung . . . . .</b>	215
31.1 Optimierung der Anwendungsentwicklung für In-Memory-Datenbanken . . . . .	215
31.1.1 Verlegen der Geschäftslogik in die Datenbank . . . . .	217
31.1.2 Stored Procedures . . . . .	218
31.1.3 Beispieldatenanwendung . . . . .	219
31.2 Best Practices . . . . .	220
31.3 Selbsttest-Fragen . . . . .	221
<b>32 Datenbank-Views . . . . .</b>	223
32.1 Vorteile von Views . . . . .	223
32.2 Konzept geschichteter Views . . . . .	224
32.3 Entwicklungswerkzeuge für Views . . . . .	224
32.4 Selbsttest-Fragen . . . . .	225
Literaturhinweise . . . . .	226
<b>33 Umgang mit Business-Objekten . . . . .</b>	227
33.1 Abbilden von Business-Objekten . . . . .	227
33.2 Objektrelationale Mapping . . . . .	228
33.3 Selbsttest-Fragen . . . . .	229
<b>34 Bypass-Lösung . . . . .</b>	231
34.1 Übergangsschritte im Detail . . . . .	231
34.2 Bypass-Lösung: Fazit . . . . .	235
34.3 Selbsttest-Frage . . . . .	236
<b>Lösungen zu den Selbsttest-Fragen . . . . .</b>	237
<b>Glossar . . . . .</b>	283