

Inhalt

Geleitwort	1
Vorwort	1
1 Einführung	1
1.1 Historie.....	1
1.2 Definition und Diskussion.....	2
1.3 Kategorisierung von NoSQL-Systemen	5
1.3.1 Key/Value-Systeme.....	7
1.3.2 Column-Family-Systeme.....	7
1.3.3 Document Stores	8
1.3.4 Graphdatenbanken.....	8
1.1 Fazit	9
2 NoSQL – Theoretische Grundlagen	11
2.1 Map/Reduce	12
2.1.1 Funktionale Ursprünge	12
2.1.2 Phasen und Datenfluss.....	17
2.1.3 Komponenten und Architektur	19
2.1.4 Anwendungsbereiche und Implementierungen.....	22
2.1.5 Praktisches Beispiel.....	24
2.1.6 Zusammenfassung	28
2.2 CAP und Eventually Consistent	30
2.2.1 Konsistenzmodell relationaler Datenbanken	30
2.2.2 CAP-Theorem	31
2.2.3 Alternatives Konsistenzmodell: BASE.....	33
2.3 Consistent-Hashing	36
2.4 Multiversion Concurrency Control	40
2.5 Vector Clocks.....	43
2.6 Paxos.....	47

3	Wide Column Stores.....	53
3.1	HBase.....	54
3.1.1	Überblick.....	54
3.1.2	Allgemeines	55
3.1.3	Datenmodell.....	55
3.1.4	Installation.....	57
3.1.5	CRUD-Operationen.....	58
3.1.6	Fortgeschrittene Abfragen mit Map/Reduce	63
3.1.7	Skalierung und Konfiguration.....	66
3.1.8	Replikation.....	68
3.1.9	Bewertung.....	68
3.2	Cassandra.....	69
3.2.1	Allgemeines	70
3.2.2	Datenmodell.....	71
3.2.3	Installation.....	74
3.2.4	CRUD-Operationen.....	76
3.2.5	Abfragen in Cassandra	79
3.2.6	Zugriff mit Java.....	79
3.2.7	Replikation und Skalierung.....	80
3.2.8	Bewertung.....	82
3.3	Amazon SimpleDB.....	83
3.3.1	Allgemeines	84
3.3.2	Datenmodell.....	85
3.3.3	Datensicherheit und Datenschutz	85
3.3.4	Installation.....	86
3.3.5	CRUD-Operationen.....	87
3.3.6	Replikation und Skalierung.....	97
3.3.7	Bewertung.....	97
4	Document Stores	101
4.1	CouchDB	102
4.1.1	Überblick.....	102
4.1.2	Beschreibung.....	102
4.1.3	Datenmodell.....	103
4.1.4	View-Modell	104
4.1.5	Zugriffskontrolle	104
4.1.6	Installation.....	104
4.1.7	CRUD-Operationen.....	107
4.1.8	Erstellen von Views	111
4.1.9	Replikation.....	112
4.1.10	Skalierung.....	113
4.1.11	CouchApps.....	113
4.1.12	Bewertung.....	114
4.2	MongoDB	115
4.2.1	Überblick.....	115
4.2.2	Datenmodell.....	116

4.2.3	Installation.....	117
4.2.4	CRUD-Operationen.....	119
4.2.5	Fortgeschrittene Abfragen und Map/Reduce.....	123
4.2.6	Skalierung.....	125
4.2.7	Replikation.....	127
4.2.8	Bewertung.....	128
5	Key/Value-Datenbanken.....	131
5.1	Redis.....	132
5.1.1	Überblick.....	132
5.1.2	Allgemeines.....	132
5.1.3	Installation.....	133
5.1.4	CRUD-Operationen.....	134
5.1.5	Redis Hashes.....	140
5.1.6	Zugriff aus anderen Programmiersprachen.....	140
5.1.7	Replikation und Konfiguration.....	142
5.1.8	Skalierung.....	144
5.1.9	Bewertung.....	144
5.2	Chordless.....	146
5.2.1	Überblick.....	146
5.2.2	Allgemeines.....	147
5.2.3	Installation.....	148
5.2.4	CRUD-Operationen.....	149
5.2.5	Fortgeschrittene Abfragen.....	151
5.2.6	Skalierung, Replikation und Konfiguration.....	154
5.2.7	Bewertung.....	154
5.3	Riak.....	156
5.3.1	Überblick.....	156
5.3.2	Allgemeines.....	157
5.3.3	Installation.....	159
5.3.4	CRUD-Operationen.....	160
5.3.5	Zugriff aus anderen Programmiersprachen.....	161
5.3.6	Abfragen und Links in Riak.....	162
5.3.7	Skalierung, Replikation und Konfiguration.....	164
5.3.8	Bewertung.....	166
6	Graphdatenbanken.....	169
6.1	Einführung und Hintergrund.....	171
6.1.1	Das Graphdatenmodell.....	172
6.1.2	Das Property-Graph-Datenmodell.....	173
6.1.3	Repräsentation von Graphen.....	174
6.1.4	Traversierung von Graphen.....	176
6.1.5	Skalierung mittels Replikation und Partitionierung.....	178
6.1.6	Vergleich mit anderen Datenmodellen.....	180
6.1.7	Zusammenfassung.....	182

6.2	Neo4j	184
6.2.1	Datenmodell	185
6.2.2	Installation und CRUD-Operationen	185
6.2.3	Fortgeschrittene Abfragen	189
6.2.4	Replikation	193
6.2.5	Horizontale Skalierung	194
6.2.6	Bewertung	194
6.3	sones	196
6.3.1	Überblick	196
6.3.2	Beschreibung	197
6.3.3	Datenmodell	198
6.3.4	Installation	199
6.3.5	CRUD-Operationen	200
6.3.6	Bewertung	208
6.4	InfoGrid	209
6.4.1	Überblick	209
6.4.2	Beschreibung	209
6.4.3	Datenmodell	210
6.4.4	Installation	211
6.4.5	CRUD-Operationen	211
6.4.6	Models	213
6.4.7	Bewertung	214
6.5	DEX	215
6.5.1	Überblick	215
6.5.2	Beschreibung	215
6.5.3	Datenmodell	216
6.5.4	Installation	216
6.5.5	CRUD-Operationen	217
6.5.6	Graphoperationen	218
6.5.7	DEX-Skripting	219
6.5.8	DEX-Shell	220
6.5.9	Bewertung	221
6.6	HyperGraphDB	222
6.6.1	Überblick	222
6.6.2	Datenmodell	223
6.6.3	Installation	224
6.6.4	CRUD-Operationen	224
6.6.5	Graphoperationen	227
6.6.6	Peer-To-Peer-Framework	227
6.6.7	Bewertung	227
6.7	InfiniteGraph	229
6.7.1	Überblick	229
6.7.2	Beschreibung	229
6.7.3	Datenmodell	230
6.7.4	Installation	230
6.7.5	CRUD-Operationen	231

6.7.6	Graphoperationen	234
6.7.7	Bewertung	235
6.8	OrientDB	236
6.8.1	Überblick	236
6.8.2	Datenmodell	237
6.8.3	Installation	237
6.8.4	CRUD-Operationen	237
6.8.5	Datenbankschemata	239
6.8.6	HTTP-REST-Schnittstelle	240
6.8.7	Bewertung	241
6.9	Weitere graphorientierte Ansätze	242
6.9.1	Twitters FlockDB	242
6.9.2	Google Pregel	244
6.9.3	Apache Hama/Hamburg	247
6.9.4	Die VertexDB-Familie	248
6.9.5	Filament	251
7	Weitere NoSQL-Datenbanken	255
7.1	Wide Column Stores	256
7.1.1	Hypertable	256
7.1.2	Cloudera	257
7.2	Document Stores	258
7.3	Key/Value/Tupel-Stores	259
7.3.1	Amazon Dynamo	259
7.3.2	Dynomite und KAI	260
7.3.3	MEMBASE	261
7.3.4	Voldemort	262
7.3.5	Scalaris	264
7.3.6	Die Tokyo-Produktfamilie	266
7.3.7	Weitere Key/Value-Systeme	267
7.4	Google App Engine Storage	268
7.5	Weitere ‚Soft‘-NoSQL-Lösungen	269
8	Orientierung im Datenbankraum	271
8.1	Grundlegende Gedanken	272
8.2	Datenanalyse	273
8.3	Transaktionsmodell	276
8.4	Performancesapekte	277
8.5	Abfrageanforderungen	277
8.6	Architektur	278
8.7	Weitere nicht-funktionale Anforderungen	279
8.8	Anwendungsfälle für NoSQL-Datenbanken	281
8.9	Fazit	282
	Register	285