

Einleitung	XIII
-------------------------	-------------

Teil I: Grundlagen von Datenystemen

1	Zuverlässige, skalierbare und wartbare Anwendungen	3
	Gedanken zu Datenystemen	4
	Zuverlässigkeit	6
	Hardwarefehler	7
	Softwarefehler	9
	Menschliche Fehler	10
	Wie wichtig ist Zuverlässigkeit?	11
	Skalierbarkeit	11
	Lasten beschreiben	11
	Performance beschreiben	14
	Konzepte zur Bewältigung von Belastungen	18
	Wartbarkeit	20
	Betriebsfähigkeit: Den Betrieb erleichtern	21
	Einfachheit: Komplexität im Griff	22
	Evolvierbarkeit: Änderungen erleichtern	23
	Zusammenfassung	24
2	Datenmodelle und Abfragesprachen	29
	Relationales Modell vs. Dokumentmodell	30
	Die Geburt von NoSQL	31
	Die objektrelationale Unverträglichkeit	32
	n:1- und n:n-Beziehungen	35
	Wiederholen Dokumentdatenbanken die Geschichte?	38
	Heutige relationale Datenbanken vs. Dokumentdatenbanken	41

Abfragesprachen für Daten	45
Deklarative Abfragen im Web	47
MapReduce-Abfragen	49
Graphen-ähnliche Datenmodelle	52
Property-Graphen	53
Die Abfragesprache Cypher	55
Graph-Abfragen in SQL	57
Triple-Stores und SPARQL	59
Das Fundament: Datalog	63
Zusammenfassung	66
3 Speichern und Abrufen	73
Datenstrukturen, auf denen Ihre Datenbank beruht	74
Hash-Indizes	76
SSTables und LSM-Bäume	80
B-Bäume	85
B-Bäume und LSM-Bäume im Vergleich	89
Andere Indizierungsstrukturen	91
Transaktionsverarbeitung oder Datenanalyse?	97
Data-Warehousing	98
Sterne und Schneeflocken: Schemas für die Analytik	101
Spaltenorientierte Speicherung	103
Spaltenkomprimierung	105
Sortierreihenfolge in spaltenorientierten Datenbanken	107
In spaltenorientierte Datenbanken schreiben	108
Aggregation: Datenwürfel und materialisierte Sichten	109
Zusammenfassung	111
4 Codierung und Evolution	119
Formate für das Codieren von Daten	120
Sprachspezifische Formate	121
JSON, XML und binäre Varianten	122
Thrift und Protocol Buffers	125
Avro	130
Die Vorzüge von Schemas	136
Datenflussmodi	137
Datenfluss über Datenbanken	137
Datenfluss über Dienste: REST und RPC	140
Datenfluss beim Nachrichtenaustausch	146
Zusammenfassung	148

Teil II: Verteilte Daten

5	Replikation	159
	Leader und Follower	160
	Synchrone und asynchrone Replikation	161
	Neue Follower einrichten	163
	Knotenausfälle behandeln	164
	Implementierung von Replikationsprotokollen	167
	Probleme mit der Replikationsverzögerung	170
	Die eigenen Schreiboperationen lesen	171
	Monotones Lesen	173
	Präfixkonsistenz	175
	Lösungen für Replikationsverzögerung	176
	Multi-Leader-Replikation	177
	Einsatzfälle für Multi-Leader-Replikation	177
	Schreibkonflikte behandeln	180
	Topologien für Multi-Leader-Replikation	185
	Replikation ohne Leader	187
	In die Datenbank schreiben, wenn ein Knoten ausgefallen ist	188
	Grenzen der Quorumkonsistenz	191
	Sloppy Quoren und Hinted Handoff	194
	Parallele Schreibvorgänge erkennen	196
	Zusammenfassung	204
6	Partitionierung	211
	Partitionierung und Replikation	212
	Partitionierung von Schlüssel-Wert-Daten	213
	Partitionierung nach Schlüsselbereich	214
	Nach dem Hashwert des Schlüssels partitionieren	215
	Schiefe Arbeitslasten und Entlastung von Hotspots	217
	Partitionierung und Sekundärindizes	218
	Sekundärindizes nach Dokument partitionieren	219
	Sekundärindizes nach Begriff partitionieren	220
	Rebalancing – Partitionen gleichmäßig belasten	222
	Strategien für Rebalancing	222
	Operationen: Automatisches oder manuelles Rebalancing	226
	Anfragen weiterleiten	227
	Parallele Abfrageausführung	229
	Zusammenfassung	230

7	Transaktionen	235
	Das schwammige Konzept einer Transaktion	236
	Die Bedeutung von ACID	237
	Einzelobjekt- und Multiobjektoperationen	242
	Schwache Isolationsstufen	248
	Read Committed	249
	Snapshot-Isolation und Repeatable Read	252
	Verlorene Updates verhindern	258
	Schreibversatz und Phantome	262
	Serialisierbarkeit	268
	Tatsächliche serielle Ausführung	269
	Zwei-Phasen-Sperrverfahren (2PL)	274
	Serialisierbare Snapshot-Isolation (SSI)	278
	Zusammenfassung	284
8	Die Probleme mit verteilten Systemen	291
	Fehler und Teilausfälle	292
	Cloud-Computing und Supercomputing	293
	Unzuverlässige Netzwerke	296
	Netzwerkfehler in der Praxis	298
	Fehler erkennen	299
	Timeouts und unbeschränkte Verzögerungen	300
	Synchrone und asynchrone Netzwerke	304
	Unzuverlässige Uhren	307
	Monotone Uhren und Echtzeituhren	308
	Uhrensynchronisierung und Genauigkeit	309
	Sich auf synchronisierte Uhren verlassen	311
	Prozesspausen	316
	Wissen, Wahrheit und Lügen	321
	Die Wahrheit wird von der Mehrheit definiert	322
	Byzantinische Fehler	326
	Systemmodell und Realität	329
	Zusammenfassung	333
9	Konsistenz und Konsens	343
	Konsistenzgarantien	344
	Linearisierbarkeit	346
	Was macht ein System linearisierbar?	347
	Auf Linearisierbarkeit setzen	352
	Linearisierbare Systeme implementieren	355
	Die Kosten der Linearisierbarkeit	358

Ordnungsgarantien	362
Ordnung und Kausalität	363
Ordnung nach Sequenznummern	368
Total geordneter Broadcast	372
Verteilte Transaktionen und Konsens	377
Atomarer Commit und Zwei-Phasen-Commit (2PC)	379
Verteilte Transaktionen in der Praxis	386
Fehlertoleranter Konsens	391
Mitgliedschafts- und Koordinationsdienste	397
Zusammenfassung	401

Teil III: Abgeleitete Daten

10 Stapelverarbeitung	417
Stapelverarbeitung mit Unix-Tools	419
Einfache Protokollanalyse	419
Die Unix-Philosophie	422
MapReduce und verteilte Dateisysteme	426
MapReduce-Jobausführung	428
Reduce-seitige Verknüpfungen und Gruppierungen	432
Map-seitige Verknüpfungen	438
Die Ausgabe von Stapel-Workflows	441
Hadoop im Vergleich mit verteilten Datenbanken	446
Jenseits von MapReduce	450
Zwischenzustände materialisieren	451
Graphen und iterative Verarbeitung	456
Höhere APIs und Sprachen	460
Zusammenfassung	462
11 Stream-Verarbeitung	471
Ereignisströme übertragen	472
Nachrichtensysteme	473
Partitionierte Protokolle	479
Datenbanken und Streams	485
Systeme synchron halten	486
Erfassen von Datenänderungen	487
Event Sourcing	491
Zustand, Streams und Unveränderlichkeit	494

Streams verarbeiten	499
Anwendungen der Stream-Verarbeitung	501
Überlegungen zur Zeit	505
Stream-Joins	509
Fehlertoleranz	513
Zusammenfassung	517
12 Die Zukunft von Datenystemen	527
Datenintegration	528
Spezialisierte Tools durch Ableiten von Daten kombinieren	528
Batch- und Stream-Verarbeitung	533
Die Entflechtung von Datenbanken	538
Zusammenstellung verschiedener Datenspeichertechniken	539
Anwendungen datenflussorientiert entwickeln	544
Abgeleitete Zustände beobachten	550
Auf der Suche nach Korrektheit	557
Das Ende-zu-Ende-Argument für Datenbanken	558
Durchsetzung von Einschränkungen	564
Zeitnähe und Integrität	567
Vertrauen ist gut, Kontrolle ist besser	572
Das Richtige tun	578
Prädiktive Analytik	579
Datenschutz und Nachverfolgung	582
Zusammenfassung	590
13 Glossar	601
Index	609