

# Contents

Preface .....	xiii
Part I: Understand .....	1
Chapter 1: Why NoSQL? .....	3
1.1 The Value of Relational Databases .....	3
1.1.1 <i>Getting at Persistent Data</i> .....	3
1.1.2 <i>Concurrency</i> .....	4
1.1.3 <i>Integration</i> .....	4
1.1.4 <i>A (Mostly) Standard Model</i> .....	4
1.2 Impedance Mismatch .....	5
1.3 Application and Integration Databases .....	6
1.4 Attack of the Clusters .....	8
1.5 The Emergence of NoSQL .....	9
1.6 Key Points .....	12
Chapter 2: Aggregate Data Models .....	13
2.1 Aggregates .....	14
2.1.1 <i>Example of Relations and Aggregates</i> .....	14
2.1.2 <i>Consequences of Aggregate Orientation</i> .....	19
2.2 Key-Value and Document Data Models .....	20
2.3 Column-Family Stores .....	21
2.4 Summarizing Aggregate-Oriented Databases .....	23
2.5 Further Reading .....	24
2.6 Key Points .....	24
Chapter 3: More Details on Data Models .....	25
3.1 Relationships .....	25
3.2 Graph Databases .....	26

3.3 Schemaless Databases .....	28
3.4 Materialized Views .....	30
3.5 Modeling for Data Access .....	31
3.6 Key Points .....	36
<b>Chapter 4: Distribution Models .....</b>	<b>37</b>
4.1 Single Server .....	37
4.2 Sharding .....	38
4.3 Master-Slave Replication .....	40
4.4 Peer-to-Peer Replication .....	42
4.5 Combining Sharding and Replication .....	43
4.6 Key Points .....	44
<b>Chapter 5: Consistency .....</b>	<b>47</b>
5.1 Update Consistency .....	47
5.2 Read Consistency .....	49
5.3 Relaxing Consistency .....	52
5.3.1 <i>The CAP Theorem</i> .....	53
5.4 Relaxing Durability .....	56
5.5 Quorums .....	57
5.6 Further Reading .....	59
5.7 Key Points .....	59
<b>Chapter 6: Version Stamps .....</b>	<b>61</b>
6.1 Business and System Transactions .....	61
6.2 Version Stamps on Multiple Nodes .....	63
6.3 Key Points .....	65
<b>Chapter 7: Map-Reduce .....</b>	<b>67</b>
7.1 Basic Map-Reduce .....	68
7.2 Partitioning and Combining .....	69
7.3 Composing Map-Reduce Calculations .....	72
7.3.1 <i>A Two Stage Map-Reduce Example</i> .....	73
7.3.2 <i>Incremental Map-Reduce</i> .....	76
7.4 Further Reading .....	77
7.5 Key Points .....	77
<b>Part II: Implement .....</b>	<b>79</b>
<b>Chapter 8: Key-Value Databases .....</b>	<b>81</b>
8.1 What Is a Key-Value Store .....	81
8.2 Key-Value Store Features .....	83

8.2.1	<i>Consistency</i>	83
8.2.2	<i>Transactions</i>	84
8.2.3	<i>Query Features</i>	84
8.2.4	<i>Structure of Data</i>	86
8.2.5	<i>Scaling</i>	86
8.3	<i>Suitable Use Cases</i>	87
8.3.1	<i>Storing Session Information</i>	87
8.3.2	<i>User Profiles, Preferences</i>	87
8.3.3	<i>Shopping Cart Data</i>	87
8.4	<i>When Not to Use</i>	87
8.4.1	<i>Relationships among Data</i>	87
8.4.2	<i>Multioperation Transactions</i>	88
8.4.3	<i>Query by Data</i>	88
8.4.4	<i>Operations by Sets</i>	88
<b>Chapter 9:</b>	<b>Document Databases</b>	<b>89</b>
9.1	<i>What Is a Document Database?</i>	90
9.2	<i>Features</i>	91
9.2.1	<i>Consistency</i>	91
9.2.2	<i>Transactions</i>	92
9.2.3	<i>Availability</i>	93
9.2.4	<i>Query Features</i>	94
9.2.5	<i>Scaling</i>	95
9.3	<i>Suitable Use Cases</i>	97
9.3.1	<i>Event Logging</i>	97
9.3.2	<i>Content Management Systems, Blogging Platforms</i>	98
9.3.3	<i>Web Analytics or Real-Time Analytics</i>	98
9.3.4	<i>E-Commerce Applications</i>	98
9.4	<i>When Not to Use</i>	98
9.4.1	<i>Complex Transactions Spanning Different Operations</i>	98
9.4.2	<i>Queries against Varying Aggregate Structure</i>	98
<b>Chapter 10:</b>	<b>Column-Family Stores</b>	<b>99</b>
10.1	<i>What Is a Column-Family Data Store?</i>	99
10.2	<i>Features</i>	100
10.2.1	<i>Consistency</i>	103
10.2.2	<i>Transactions</i>	104
10.2.3	<i>Availability</i>	104



10.2.4	<i>Query Features</i> .....	105
10.2.5	<i>Scaling</i> .....	107
10.3	<i>Suitable Use Cases</i> .....	107
10.3.1	<i>Event Logging</i> .....	107
10.3.2	<i>Content Management Systems, Blogging Platforms</i> .....	108
10.3.3	<i>Counters</i> .....	108
10.3.4	<i>Expiring Usage</i> .....	108
10.4	<i>When Not to Use</i> .....	109
<b>Chapter 11:</b>	<b>Graph Databases</b> .....	<b>111</b>
11.1	<i>What Is a Graph Database?</i> .....	111
11.2	<i>Features</i> .....	113
11.2.1	<i>Consistency</i> .....	114
11.2.2	<i>Transactions</i> .....	114
11.2.3	<i>Availability</i> .....	115
11.2.4	<i>Query Features</i> .....	115
11.2.5	<i>Scaling</i> .....	119
11.3	<i>Suitable Use Cases</i> .....	120
11.3.1	<i>Connected Data</i> .....	120
11.3.2	<i>Routing, Dispatch, and Location-Based Services</i> .....	120
11.3.3	<i>Recommendation Engines</i> .....	121
11.4	<i>When Not to Use</i> .....	121
<b>Chapter 12:</b>	<b>Schema Migrations</b> .....	<b>123</b>
12.1	<i>Schema Changes</i> .....	123
12.2	<i>Schema Changes in RDBMS</i> .....	123
12.2.1	<i>Migrations for Green Field Projects</i> .....	124
12.2.2	<i>Migrations in Legacy Projects</i> .....	126
12.3	<i>Schema Changes in a NoSQL Data Store</i> .....	128
12.3.1	<i>Incremental Migration</i> .....	130
12.3.2	<i>Migrations in Graph Databases</i> .....	131
12.3.3	<i>Changing Aggregate Structure</i> .....	132
12.4	<i>Further Reading</i> .....	132
12.5	<i>Key Points</i> .....	132
<b>Chapter 13:</b>	<b>Polyglot Persistence</b> .....	<b>133</b>
13.1	<i>Disparate Data Storage Needs</i> .....	133
13.2	<i>Polyglot Data Store Usage</i> .....	134
13.3	<i>Service Usage over Direct Data Store Usage</i> .....	136

13.4	Expanding for Better Functionality .....	136
13.5	Choosing the Right Technology .....	138
13.6	Enterprise Concerns with Polyglot Persistence .....	138
13.7	Deployment Complexity .....	139
13.8	Key Points .....	140
<b>Chapter 14:</b>	<b>Beyond NoSQL .....</b>	<b>141</b>
14.1	File Systems .....	141
14.2	Event Sourcing .....	142
14.3	Memory Image .....	144
14.4	Version Control .....	145
14.5	XML Databases .....	145
14.6	Object Databases .....	146
14.7	Key Points .....	146
<b>Chapter 15:</b>	<b>Choosing Your Database .....</b>	<b>147</b>
15.1	Programmer Productivity .....	147
15.2	Data-Access Performance .....	149
15.3	Sticking with the Default .....	150
15.4	Hedging Your Bets .....	150
15.5	Key Points .....	151
15.6	Final Thoughts .....	152
<b>Bibliography</b>	<b>.....</b>	<b>153</b>
<b>Index</b>	<b>.....</b>	<b>157</b>