

《解读NoSQL》

图书基本信息

书名：《解读NoSQL》

13位ISBN编号：9787115411107

作者：[美] 丹·麦克雷 (Dan McCreary), [美] 安·凯利 (Ann Kelly)

页数：237

译者：范东来,滕雨橦

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：www.tushu000.com

内容概要

本书从NoSQL的相关理论开始，深入浅出地探讨了NoSQL核心的架构模式、解决方案和一些高级主题，内容循序渐进，从理论回归于实践。全书分为4个部分。首部分介绍NoSQL的相关理论，如CAP理论、BASE理论、一致性散列算法等；第二部分介绍NoSQL核心的架构模式—键值存储、图存储、列族存储、文档存储；第三部分展现一些常用的NoSQL解决方案，如HA、全文搜索等；后一部分讨论NoSQL的一些高级主题，如函数式编程。全书理论与实践并重，每章后面还有通俗的案例。对于NoSQL的初学者来说，不失为一本了解NoSQL技术全貌的读物。

作者简介

Dan McCreary 和Ann Kelly领导了一家独立的培训与咨询机构，专注于NoSQL解决方案，并且他们还是NoSQL Now！会议的联合发起人。

书籍目录

第一部分 了解NoSQL

第1章 NoSQL：明智的选择 2

1.1 什么是NoSQL 3

1.2 NoSQL的商业驱动 4

1.2.1 容量 5

1.2.2 速度 5

1.2.3 敏捷性 6

1.3 NoSQL案例研究 6

1.3.1 案例研究：LiveJournal的Memcache技术 7

1.3.2 案例研究：Google的MapReduce——利用商用硬件生成搜索索引 7

1.3.3 案例研究：Google的Bigtable——一个有着数十亿行和百万列的表 8

1.3.4 案例研究：亚马逊的Dynamo——每天24小时接收订单 9

1.3.5 案例研究：MarkLogic 9

1.3.6 实践 10

1.4 小结 10

第2章 NoSQL概念 12

2.1 保持组件简单以促进重用 12

2.2 将应用分层以简化设计 14

2.3 策略地使用RAM、SSD和磁盘提升性能 17

2.4 使用一致性散列算法维护当前的缓存 18

2.5 比较ACID和BASE——两种可靠的数据库事务方法 19

2.5.1 RDBMS的事务控制——ACID 21

2.5.2 非RDBMS的事务控制——BASE 22

2.6 通过数据库分片获得水平扩展能力 23

2.7 基于Brewer的CAP定理进行权衡 25

2.8 实践 26

2.9 小结 27

2.10 延伸阅读 27

第二部分 数据库模式

第3章 基础数据架构模式 30

3.1 什么是数据架构模式 31

3.2 理解应用于RDBMS的行存储设计模式 31

3.2.1 行存储如何工作 32

3.2.2 行存储的演变 33

3.2.3 分析行存储模式的优点和缺点 34

3.3 示例：对销售订单进行连接操作 35

3.4 回顾RDBMS实现的特性 36

3.4.1 RDBMS事务 37

3.4.2 固定的数据定义语言和强类型的列 38

3.4.3 通过RDBMS视图保证安全并进行访问控制 39

3.4.4 RDBMS的复制和同步 39

3.5 通过OLAP、数据仓库和商业智能系统对历史数据进行分析 41

3.5.1 数据如何从操作型系统流入分析型系统 42

3.5.2 熟悉OLAP的概念 44

3.5.3 通过汇总生成专项报表 45

3.6 将高可用性和以读为主的系统一体化 46

3.7 在修改控制系统和数据同步中使用散列树 47

- 3.8 实践 49
- 3.9 小结 49
- 3.10 延伸阅读 50
- 第4章 NoSQL数据架构模式 51
 - 4.1 键值存储 52
 - 4.1.1 什么是键值存储 52
 - 4.1.2 使用键值存储的好处 54
 - 4.1.3 使用键值存储 56
 - 4.1.4 使用案例：用键值存储保存网页 59
 - 4.1.5 使用案例：亚马逊简单存储服务（S3） 59
 - 4.2 图存储 60
 - 4.2.1 图存储概述 60
 - 4.2.2 用RDF标准来连接外部数据 62
 - 4.2.3 图存储的使用案例 63
 - 4.3 列族（Bigtable）存储 68
 - 4.3.1 列族存储基础 69
 - 4.3.2 理解列族存储的键 69
 - 4.3.3 列族存储的优点 70
 - 4.3.4 案例研究：在Bigtable中存储分析信息 72
 - 4.3.5 案例研究：Google地图用Bigtable存储地理信息 72
 - 4.3.6 案例研究：使用列族存储用户偏好信息 73
 - 4.4 文档存储 73
 - 4.4.1 文档存储基础 74
 - 4.4.2 文档集合 74
 - 4.4.3 应用集合 75
 - 4.4.4 文档存储的API 75
 - 4.4.5 文档存储的实现 76
 - 4.4.6 案例研究：MongoDB和广告服务器 76
 - 4.4.7 案例研究：大型对象数据库CouchDB 77
 - 4.5 NoSQL架构模式的变体 78
 - 4.5.1 定制RAM和SSD存储 78
 - 4.5.2 分布式存储 78
 - 4.5.3 分组的对象 79
 - 4.6 小结 80
 - 4.7 延伸阅读 81
- 第5章 原生XML数据库 82
 - 5.1 什么是原生XML数据库 83
 - 5.2 用原生XML数据库构建应用 85
 - 5.2.1 加载数据可以像拖曳那样简单 86
 - 5.2.2 使用集合来组织XML文档 87
 - 5.2.3 使用XPath运用简单的查询转换复杂的数据 89
 - 5.2.4 用XQuery转换数据 91
 - 5.2.5 用XQuery更新文档 93
 - 5.2.6 XQuery全文搜索标准 94
 - 5.3 在原生XML数据库中应用XML标准 94
 - 5.4 用XML Schema和Schematron设计和验证数据 96
 - 5.4.1 XML Schema 96
 - 5.4.2 使用Schematron检查文档规则 97
 - 5.5 用自定义模块扩展XQuery 98

- 5.6 案例研究：在美国国务院历史学家办公室使用NoSQL 98
- 5.7 案例研究：使用MarkLogic管理金融衍生品 102
 - 5.7.1 为什么RDBMS难以存储金融衍生品 102
 - 5.7.2 一个投资银行从20个RDBMS转换到1个原生XML数据库 102
 - 5.7.3 迁移至原生XML文档存储的商业好处 104
 - 5.7.4 项目成果 104
- 5.8 小结 105
- 5.9 延伸阅读 105
- 第三部分 NoSQL解决方案
- 第6章 用NoSQL管理大数据 108
 - 6.1 什么才是大数据解决方案 109
 - 6.2 线性扩展数据中心 112
 - 6.3 理解线性可扩展性和表现力 113
 - 6.4 了解大数据问题的类型 114
 - 6.5 使用无共享架构分析大数据 116
 - 6.6 选择分布式模型：主从模型与对等模型 117
 - 6.7 在分布式系统上使用MapReduce处理数据 118
 - 6.7.1 MapReduce和分布式文件系统 120
 - 6.7.2 MapReduce怎样做到高效处理大数据问题 121
 - 6.8 NoSQL系统处理大数据问题的4种方式 122
 - 6.8.1 分发查询到数据，而非数据到查询 122
 - 6.8.2 使用散列环在集群中均匀分发数据 122
 - 6.8.3 使用复制扩展读取性能 123
 - 6.8.4 使数据库将查询均衡地分发到数据节点 124
 - 6.9 案例研究：使用ApacheFlume处理事件日志 125
 - 6.9.1 事件日志数据分析的挑战 125
 - 6.9.2 Apache Flume搜集分布式事件日志的方法 126
 - 6.9.3 延伸思考 127
 - 6.10 案例研究：计算机辅助发现医疗保险欺诈 128
 - 6.10.1 什么是医疗保险欺诈检测 128
 - 6.10.2 使用图和定制的内网共享硬件检测医疗保险欺诈 129
 - 6.11 小结 130
 - 6.12 延伸阅读 131
- 第7章 用NoSQL搜索获取信息 132
 - 7.1 什么是NoSQL搜索 132
 - 7.2 搜索分类 133
 - 7.2.1 布尔搜索、全文关键字搜索和结构化搜索的比较 133
 - 7.2.2 测试常见搜索类型 134
 - 7.3 提高NoSQL搜索效率的策略和方法 135
 - 7.4 使用文档结构提升搜索质量 137
 - 7.5 搜索质量量化 139
 - 7.6 本地索引与远程搜索服务 139
 - 7.7 案例研究：使用MapReduce建立倒排索引 141
 - 7.8 案例研究：搜索技术文档 142
 - 7.8.1 什么是技术文档搜索 142
 - 7.8.2 在NoSQL文档存储中保留文档结构 143
 - 7.9 案例研究：搜索领域语言—可检索性和重用性 144
 - 7.10 实践 145
 - 7.11 小结 146

- 7.12 延伸阅读 146
- 第8章 用NoSQL构建高可用的解决方案 148
 - 8.1 高可用NoSQL数据库的定义 148
 - 8.2 量化NoSQL数据库的可用性 149
 - 8.2.1 案例研究：亚马逊S3的服务级别协议 151
 - 8.2.2 预测系统可用性 151
 - 8.2.3 实践 152
 - 8.3 NoSQL系统的高可用性策略 153
 - 8.3.1 使用负载均衡器将流量转向到最空闲的节点 153
 - 8.3.2 结合高可用分布式文件系统和NoSQL数据库 154
 - 8.3.3 案例研究：将HDFS作为一个高可用的文件系统存储主数据 155
 - 8.3.4 使用托管的NoSQL服务 156
 - 8.3.5 案例研究：使用亚马逊的DynamoDB作为高可用数据存储 157
 - 8.4 案例研究：使用ApacheCassandra作为高可用的列族存储 158
 - 8.4.1 在Cassandra中配置数据和节点间的映射 159
 - 8.5 案例研究：使用Couchbase作为高可用文档数据库 161
 - 8.6 小结 163
 - 8.7 延伸阅读 163
- 第9章 用NoSQL提升敏捷性 165
 - 9.1 软件敏捷性的定义 165
 - 9.2 量化敏捷性 169
 - 9.3 使用文档存储来避免对象关系映射 171
 - 9.4 案例研究：使用XRX管理复杂表单 172
 - 9.4.1 什么是复杂业务表单 173
 - 9.4.2 用XRX替换客户端JavaScript和对象关系映射 173
 - 9.4.3 理解XRX对敏捷性的影响 176
 - 9.5 小结 177
 - 9.6 延伸阅读 177
- 第四部分 高级主题
- 第10章 NoSQL与函数式编程 180
 - 10.1 什么是函数式编程 181
 - 10.1.1 指令式编程就是管理程序状态 181
 - 10.1.2 函数式编程是没有副作用的并行转化 183
 - 10.1.3 比较指令式编程和函数式编程的扩展性 186
 - 10.1.4 使用引用透明避免重复计算 187
 - 10.2 案例研究：用NetKernel优化网页内容组装 188
 - 10.2.1 组装嵌套内容，追踪组件依赖 188
 - 10.2.2 用NetKernel优化组件再生成 189
 - 10.3 函数式编程语言示例 191
 - 10.4 完成指令式编程到函数式的编程转变 192
 - 10.4.1 使用函数作为函数的参数 192
 - 10.4.2 使用递归处理非结构化文档数据 192
 - 10.4.3 使用不可变变量而非可变变量 192
 - 10.4.4 去除循环和条件语句 193
 - 10.4.5 新的思维方式：从状态记录到转化隔离 193
 - 10.4.6 质量、校验和一致性单元测试 194
 - 10.4.7 函数式编程的并发 194
 - 10.5 案例研究：用Erlang构建NoSQL系统 194
 - 10.6 实践 197

10.7	小结	198
10.8	延伸阅读	198
第11章	安全：保护NoSQL系统中的数据	200
11.1	NoSQL数据库的一种安全模型	201
11.1.1	使用服务减少数据库内部的安全性需求	202
11.1.2	使用数据仓库和OLAP减少数据库内部的安全性需求	203
11.1.3	应用级安全措施和数据库级安全措施的收益总结	203
11.2	收集安全需求	204
11.2.1	认证	205
11.2.2	授权	207
11.2.3	审查和日志记录	210
11.2.4	加密和数字签名	211
11.2.5	保护公开网站免受拒绝服务攻击和注入攻击	212
11.3	案例研究：键值存储的访问控制—亚马逊的S3	213
11.3.1	身份和访问管理（IAM）	214
11.3.2	访问控制列表（ACL）	214
11.3.3	桶策略	214
11.4	案例研究：在Apache Accumulo中使用键可见性技术	215
11.5	案例研究：在安全发布流程中使用MarkLogic的RBAC模型	216
11.5.1	使用MarkLogic的RBAC安全模型保护文档	217
11.5.2	在安全出版中使用MarkLogic	218
11.5.3	MarkLogic的安全模型的优势	218
11.6	小结	219
11.7	延伸阅读	219
第12章	选择合适的NoSQL解决方案	221
12.1	什么是架构利弊分析	222
12.2	数据库架构选型团队的组成变化	223
12.2.1	选择合适的团队	224
12.2.2	考虑经验偏好	224
12.2.3	雇用外部咨询师	225
12.3	架构权衡分析步骤	225
12.4	解构分析：质量树	228
12.4.1	质量属性样例	229
12.4.2	评估混合架构和云架构	231
12.5	与利益系相关者沟通结果	231
12.5.1	用质量树作为导航图	232
12.5.2	实践	233
12.5.3	使用质量树进行项目风险交流	234
12.6	找到合适的验证架构的试点项目	235
12.7	小结	236
12.8	延伸阅读	237

精彩短评

- 1、这本书貌似没什么人读，其实还不错
- 2、理解NoSQL一个很合适的口子。

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:www.tushu000.com