

Cassandra实战



[Cassandra实战_下载链接1](#)

著者:郭鹏

出版者:机械工业

出版时间:2011-6

装帧:

isbn:9787111341642

《Cassandra实战》是国内首本Cassandra专著，由Cassandra领域的先驱者和实践者亲自执笔，多位数据库专家联袂推荐，权威性毋庸置疑。

本书内容全面，基于Cassandra最新版撰写，系统地讲解了Cassandra的所有功能特性和使用方法；实战性强，不仅包含大量示例代码，而且还设计了一个完整的在线交易系统实例；有一定的深度，不仅结合源代码分析了Cassandra的底层机制和工作原理，而且还精心总结了一些关于Cassandra的最佳实践。

本书一共分为13章，首先简单介绍了NoSQL的优势，以及几种具有代表性的NoSQL数据库的功能特性；其次详细讲解了Cassandra的安装和配置、数据模型和排序规则、编程接口等基础知识；接着以迭代的方式演示了一个基于Cassandra的在线交易系统的完整开发过程，很好地将基础理论融入到了实践中；再接着结合源代码分析了Cassandra的集群机制、内部数据存储结构、数据更新机制、数据读取机制、数据压缩机制、启动流程等与Cassandra的底层机制和工作原理相关的内容；最后讲解了Cassandra在分布式环境中的应用、与Hadoop的整合，以及相关的最佳实践。附录中包含了本书示例的源代码以及在Eclipse环境中编辑和修改Cassandra的源代码方法。

本书适合所有对Cassandra感兴趣的读者阅读。通过本书，不仅能全面掌握Cassandra的基础知识和使用方法，还能深入理解Cassandra的底层机制和工作原理，以及它在复杂现实环境中的应用。

作者介绍:

郭鹏，国内Cassandra领域的先驱者和实践者。资深软件开发工程师，擅长分布式应用程序的开发和使用，时间经验极其丰富。一直致力于Cassandra、Hadoop、Hive、MongoDB等分布式应用的底层实现原理的研究和探索，对分布式数据库和数据仓库也有深刻而独到的理解。曾任阿里巴巴数据仓库开发工程师，参与设计和开发了多个基于Cassandra和Hadoop的大型应用。现供职于盛大在线，参与Hadoop计算存储平台的搭建与实施。活跃于开源技术社区，是博客园的推荐博客，也是CSDN的博客专家，热衷于与网友分享自己的心得和体会，深受社区欢迎。博客地址：<http://gpcuster.cnblogs.com/>；新浪微博：@逖靖寒。

目录: 前言

第1章认识NoSQL/1

1.1 NoSQL的起源和发展现状/2

1.2 为什么要使用NoSQL/2

1.3 开源NoSQL产品介绍/3

1.3.1 Key/Value的NoSQL数据库/3

1.3.2 面向文档的NoSQL数据库/4

1.3.3 面向列的NoSQL数据库/5

1.3.4 面向图的NoSQL数据库/6

1.4 本章小结/7

第2章Cassandra快速入门/9

2.1 在Windows环境运行单机版Cassandra/10

2.1.1 配置JRE/11

2.1.2 配置运行Cassandra 0.6.x/11

2.1.3 配置运行Cassandra 0.7.x/12

2.2 在Linux环境运行单机版Cassandra/14

2.2.1 配置JRE/14

2.2.2 配置运行Cassandra 0.6.x/15

2.2.3 配置运行Cassandra 0.7.x/16

2.3 Cassandra的数据模型/18

2.3.1 Column/18

2.3.2 SuperColumn/18

2.3.3 ColumnFamily/19

- 2.3.4Keyspace/20
- 2.4Cassandra的数据排序规则/20
- 2.5配置数据类型/22
- 2.6使用命令行工具与Cassandra交互/23
 - 2.6.1与Cassandra 0.6.x进行交互/23
 - 2.6.2与Cassandra 0.7.x进行交互/24
- 2.7本章小结/26
- 第3章理解Cassandra编程接口/27
 - 3.1多语言服务开发框架Thrift/28
 - 3.2Cassandra的数据类型/28
 - 3.2.1Column/28
 - 3.2.2SuperColumn/29
 - 3.2.3ColumnOrSuperColumn/29
 - 3.2.4ColumnParent/29
 - 3.2.5ColumnPath/30
 - 3.2.6SliceRange/30
 - 3.2.7SlicePredicate/31
 - 3.2.8Deletion/31
 - 3.2.9Mutation/31
 - 3.2.10KeyRange/31
 - 3.2.11KeySlice/32
 - 3.2.12TokenRange/32
 - 3.2.13AuthenticationRequest/32
 - 3.2.14ConsistencyLevel/33
 - 3.2.15NotFoundException/33
 - 3.2.16InvalidRequestException/34
 - 3.2.17UnavailableException/34
 - 3.2.18TimedOutException/34
 - 3.2.19AuthenticationException/34
 - 3.2.20AuthorizationException/35
 - 3.3Cassandra的编程接口/35
 - 3.3.1get/35
 - 3.3.2get_slice/36
 - 3.3.3multiget_slice/36
 - 3.3.4get_count/37
 - 3.3.5get_range_slices/37
 - 3.3.6insert/38
 - 3.3.7remove/38
 - 3.3.8batch_mutate/39
 - 3.3.9describe_keyspaces/39
 - 3.3.10describe_keyspace/39
 - 3.3.11describe_cluster_name/40
 - 3.3.12describe_version/40
 - 3.3.13describe_ring/40
 - 3.4Cassandra 0.7.x版本新增功能/40
 - 3.4.1二级索引/40
 - 3.4.2动态修改Schema/44
 - 3.4.3自动清除过期数据/46
 - 3.5本章小结/47
- 第4章基于Cassandra的在线交易系统/49
 - 4.1需求分析/50
 - 4.2数据模型设计/50
 - 4.2.1Seller/51
 - 4.2.2Buyer/51

- 4.2.3Product/52
- 4.2.4ProductCategory/53
- 4.2.5Comment/54
- 4.3编码实现/54
 - 4.3.1修改Keyspace设置/55
 - 4.3.2建立Eclipse项目/56
 - 4.3.3实体对象实现/55
 - 4.3.4Cassandra数据操作接口实现/56
- 4.4系统功能验证/60
 - 4.4.1BuyerDao功能验证/60
 - 4.4.2SellerDao功能验证/61
 - 4.4.3ProductDao功能验证/62
- 4.5迁移到Cassandra 0.7.x/65
 - 4.5.1建立Eclipse项目/65
 - 4.5.2修改编译错误代码/65
 - 4.5.3新增Schema在线定义功能/69
 - 4.5.4功能验证/70
- 4.6本章小结/71
- 第5章Cassandra的集群机制/73
 - 5.1一致性哈希/74
 - 5.1.1理解一致性哈希/74
 - 5.1.2一致性哈希在Cassandra中的应用/77
 - 5.2Gossip：集群节点之间的通信协议/81
 - 5.2.1FailureDetector/82
 - 5.2.2Gossiper/83
 - 5.3集群的数据备份机制/88
 - 5.3.1EndpointSnitch/88
 - 5.3.2ReplicationStrategy/91
 - 5.4集群状态变化的处理机制/96
 - 5.4.1StorageLoadBalancer/97
 - 5.4.2StorageService/98
 - 5.4.3MigrationManager/99
 - 5.5本章小结/99
- 第6章Cassandra的内部数据存储结构/101
 - 6.1Cassandra中的数据存储规则/102
 - 6.2Commilog/102
 - 6.3Memtable/103
 - 6.4SSTable/105
 - 6.4.1Filter文件/105
 - 6.4.2Index文件/107
 - 6.4.3Data文件/109
 - 6.4.4Statistics文件/113
 - 6.5系统表空间/113
 - 6.6本章小结/114
- 第7章Cassandra的数据更新机制/115
 - 7.1数据更新流程/116
 - 7.2集群数据更新策略/116
 - 7.2.1ANY/120
 - 7.2.2ONE/121
 - 7.2.3QUORUM/121
 - 7.2.4LOCAL_QUORUM/121
 - 7.2.5EACH_QUORUM/121
 - 7.2.6ALL/121
 - 7.3二级索引/122

- 7.3.1为什么需要二级索引/122
- 7.3.2Cassandra二级索引更新过程/123
- 7.4本章小结/124
- 第8章Cassandra的数据读取机制/125
- 8.1数据读取流程/126
 - 8.1.1弱读取/126
 - 8.1.2强读取/128
- 8.2集群数据读取策略/131
 - 8.2.1ONE/132
 - 8.2.2QUORUM/132
 - 8.2.3LOCAL_QUORUM/132
 - 8.2.4EACH_QUORUM/132
 - 8.2.5ALL/133
- 8.3读修复/133
- 8.4数据缓存/134
 - 8.4.1RowCache/134
 - 8.4.2KeyCache/134
- 8.5二级索引/135
- 8.6本章小结/135
- 第9章Cassandra的数据压缩机制/137
- 9.1为什么要进行数据压缩/138
- 9.2如何控制数据压缩/138
- 9.3数据压缩流程/139
- 9.4维护Cassandra中的数据/143
 - 9.4.1数据清理压缩/143
 - 9.4.2数据一致性校验压缩/144
- 9.5本章小结/144
- 第10章Cassandra的启动流程/145
- 10.1Cassandra启动脚本/146
- 10.2Cassandra启动流程/149
 - 10.2.1配置log4j/150
 - 10.2.2读取校验配置文件信息/150
 - 10.2.3加载所有的数据文件/152
 - 10.2.4修复数据/154
 - 10.2.5启动Gossiper服务/155
 - 10.2.6判断是否需要进行Bootstrap操作/156
 - 10.2.7监听Thrift端口，提供Thrift服务/157
- 10.3本章小结/157
- 第11章在分布式环境中使用的Cassandra/159
- 11.1在Linux环境中搭建与使用Cassandra集群/160
 - 11.1.1配置JRE/160
 - 11.1.2部署Cassandra可执行文件/161
 - 11.1.3修改Cassandra配置文件/162
 - 11.1.4启动Cassandra/163
- 11.2Cassandra运行配置项详解/166
- 11.3Cassandra集群的运行和维护/175
 - 11.3.1查看集群的运行情况/176
 - 11.3.2添加节点/179
 - 11.3.3删除节点/181
 - 11.3.4移动节点/183
 - 11.3.5数据维护/185
- 11.4本章小结/187
- 第12章Cassandra与Hadoop的整合/189
- 12.1Hadoop快速入门/190

- 12.1.1Hadoop简介/190
- 12.1.2HDFS/192
- 12.1.3Map/Reduce/192
- 12.1.4配置单机版Hadoop/194
- 12.1.5编写Map/Reduce程序/195
- 12.2为什么要整合Cassandra与Hadoop/200
- 12.3使用Map/Reduce导入数据到Cassandra中/200
- 12.4将Cassandra中的数据作为Map/Reduce输入/205
- 12.5本章小结/209
- 第13章Cassandra最佳实践/211
- 13.1避免Cassandra自身的限制/212
 - 13.1.1不要盲目使用Super Column/212
 - 13.1.2硬盘的容量大小限制/212
 - 13.1.3注意系统大小限制/212
- 13.2数据压缩策略/213
- 13.3使用高级的客户端/213
 - 13.3.1Pycassa/213
 - 13.3.2Hector/215
 - 13.3.3FluentCassandra/218
 - 13.3.4Cassandra/220
 - 13.3.5phpcassa/221
- 13.4负载均衡/222
 - 13.4.1随机选取/222
 - 13.4.2缓存集群信息/222
- 13.5谨慎使用二级索引/223
- 13.6通过JMX监测Cassandra/223
- 13.7调整JVM启动参数/229
- • • • • ([收起](#))

[Cassandra实战_下载链接1](#)

标签

Cassandra

nosql

计算机

数据库

NoSQL

海量数据存储

分布式

藏书

评论

国内第一本Cassandra著作，基于最新版Cassandra撰写，杨海朝、姜承尧、孙立等数位数据库专家联袂推荐。

这本不错，适合两小时快速阅读，结合源码把大部分的点都讲到了

看了一遍，很多地方感觉讲的不是很清楚，可能我对于nosql的基础太差了。准备再去看看cassandra: the definitiv guide。

代码实在有点太多了

对cassandra的源码分析的不错，本来想看源码的，但是一拖再拖就看不了了。

包括一些概念性介绍，也分析了部分源码及实现原理，有些实战经验分享；但是坑爹是附录贴了近50页源代码，不知道用svn、git分享书中源码么，严重怀疑是凑页数的...

小众nosql，谨慎阅读

写的挺不错的

[Cassandra实战_下载链接1](#)

书评

大体上介绍了cassandra的结构以及运作方式以及一些简单的编程例子。
对于深入cassandra具有打基础的功效。
是一本入门的好书，尽管有些地方还不够详实，甚至一笔带过的。
最后，本书中的cassandra版本与最新的8系列版本略有偏差。
从8开始的版本做了一些调整，更为详细的信息可...

1 版本陈旧，现在已经有2.0.x，1.2.x，本书讲的是0.6.x
0.7.x，但是确是cassandra的资料比较少。（不止是书，网上资料跟mongoDB比也非常少）
2 讲的很浅，而且跳跃性比较强。
如果对bigtable、sstable、nosql的实现原理和概念不了解，后几章，如replication，compaction之类等...

《Cassandra实战》图书详情 <http://product.china-pub.com/198207>
《Cassandra实战》电子书在线阅读
<http://www.china-pub.com/computers/common/mianfeisd.asp?id=198207>

有一个好消息要告诉大家，台湾一家大型出版社上周来华章公司洽谈，初步评估了一下，觉得本书质量非常不错，拟引进本书的繁体版版权，这些台湾的读者也有福了。
作为国内第一本关于Cassandra的书，不仅是基于最新版本的，不仅全面讲解了Cassandra的操作和使用，更重要的是结合...

[Cassandra实战_下载链接1](#)