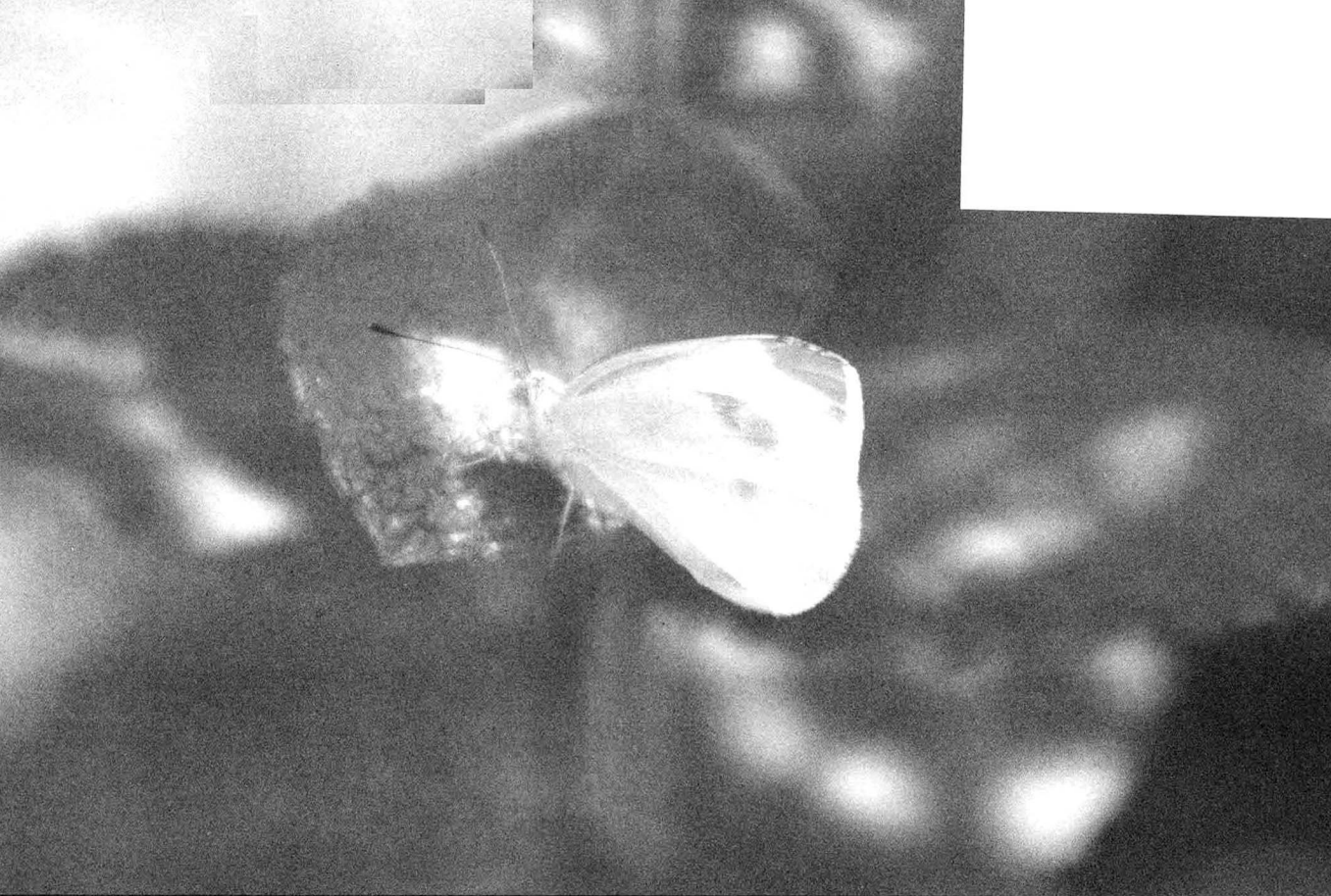


熟练掌握HBase的配置及管理技巧，最优化数据库性能

HBase管理指南

/
HBase Administration Cookbook

蒋焱峰 著
苏南 译



HBase管理指南

蒋焱峰 著
苏南 译

人民邮电出版社
北京

图书在版编目 (C I P) 数据

HBase管理指南 / 蒋燚峰著 ; 苏南译. — 北京 :
人民邮电出版社, 2013. 7
ISBN 978-7-115-31981-4

I. ①H… II. ①蒋… ②苏… III. ①计算机网络—信
息存贮—指南 IV. ①TP393-62

中国版本图书馆CIP数据核字(2013)第109603号

版 权 声 明

Copyright ©2012 Packt Publishing. First published in the English language under the title
HBase Administration Cookbook

All Rights Reserved.

本书由美国 **Packt Publishing** 公司授权人民邮电出版社出版。未经出版者书面许可, 对本书的任何部分不得以任何方式或任何手段复制和传播。

版权所有, 侵权必究。

-
- ◆ 著 蒋燚峰
 - 译 苏 南
 - 责任编辑 汪 振
 - 责任印制 程彦红 焦志炜
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
三河市海波印务有限公司印刷
 - ◆ 开本: 800×1000 1/16
印张: 19.75
字数: 389 千字 2013 年 7 月第 1 版
印数: 1-3 000 册 2013 年 7 月河北第 1 次印刷
- 著作权合同登记号 图字: 01-2012-8550 号
-

定价: 59.00 元

读者服务热线: (010)67132692 印装质量热线: (010)67129223
反盗版热线: (010)67171154

内容提要

本书通过详尽的操作步骤以及贴近实际使用的案例说明，帮助读者轻松掌握管理 HBase 所需的各项技能。本书是一部实用性很强的操作指南，主要介绍如何建立一个完全分布式的 HBase 集群并将数据转移到该集群中，以及如何使用各种工具来完成日常的管理工作，有效管理和监控集群以使其达到最佳的性能，最终还将讲解如何安装 Hadoop 集群、如何配置 Hadoop，使其与 HBase 进行协作和如何调整 Hadoop 的性能。

本书适合使用 HBase 进行数据库开发的高级数据库研发人员阅读使用。

作者简介

蒋燚峰 (Yifeng Jiang) 是日本最大的电子商务公司乐天 (Rakuten) 公司的 Hadoop 和 HBase 管理员和开发人员。他毕业于中国科学技术大学, 获得了信息管理系统专业的学士学位。在其专业软件工程师职业生涯的开始阶段, 他专注于 Java 开发。

从 2008 年开始, 他一直在寻找 Hadoop 方面的项目。2009 年, 他在上一家公司领导了一个使用 Hadoop 和 Hive 来开发的显示广告方面数据基础设施项目。

2010 年, 他加入了乐天公司, 设计并实现了一个基于 Hadoop 和 HBase 的大型商品排名系统。他还是该公司 Hadoop 团队的一员, 该团队负责多个 Hadoop/HBase 集群的运维工作。

致谢

2011 年 9 月,当 Packt 出版公司第一次问我是否有兴趣写一本关于 HBase 管理的书时,我一点也不知道这将会是一种多大的工作压力(当然,这其中也有很多的乐趣)。

现在,这本书终于完成了。我要感谢那些帮助过我的人,没有他们,本书也不可能完成。

首先,我要感谢 HBase 的开发者们,他们给我们带来了这样一个伟大的软件。感谢他们曾在邮件列表上为我的很多问题提供了很好的解答,感谢他们努力地完成了各项功能和文档。

我还要感谢 Packt 出版公司的团队,感谢他们最开始联系我写作本书,并一直对我进行支持、指导和反馈。

非常感谢乐天公司(Rakuten),我所就职的这家公司给我提供了一个使用 HBase 的工作环境,因此我才有机会写作本书。

感谢 Michael Stack, 谢谢你对本书进行的快速审校。

感谢本书的审校者——Michael Morello、河野达人(Tatsuya Kawano)、滨野贤一郎(Kenichiro Hamano)、山下真一(Shinichi Yamashita)和岩崎正刚(Masatake Iwasaki)。

感谢香川洋太郎(Yotaro Kagawa): 感谢你一直以来对于我和我的家庭的支持。

感谢信萍和铃音: 谢谢你们的支持和耐心,我爱你们!

本书审校

岩崎正刚（Masatake Iwasaki）是一名 NTT 数据公司的软件工程师，专门提供 Hadoop、HBase 以及 PostgreSQL 等开源软件方面的技术咨询。

河野达人（Tatsuya Kawano）是 HBase 在日本的贡献者和传播者。2010 年以来，他一直在帮助日本的 Hadoop 和 HBase 社区不断成长。

目前，他在为 Gemini Mobile Technologies 工作，是其研发部的一名软件工程师。他还开发了一个与 S3 API 完全兼容的云存储平台 Cloudian 和一个开源的分布式键值对存储系统 Hibari DB。

他是日本 2012 年出版的《NoSQL 基本知识》一书的合作作者，该书向初学者们介绍了包括 HBase、Cassandra、Riak、MongoDB 和 Neo4j 在内的 16 种 NoSQL 产品。

20 世纪 90 年代中后期，他曾在纽约学习平面设计。他喜欢玩 3D 电脑绘图，就如同他喜爱开发高可用、可伸缩的存储系统一样。

Michael Morello 持有分布式计算和人工智能的硕士学位。他是一名资深的高级 Java/JEE 开发人员，具有深厚的 UNIX 和 Linux 背景。他的研究领域大多与大型系统和专门解决可伸缩性、性能和高可用性问题的新兴技术有关。

我想感谢我的妻子和我的小天使，感谢她们的爱和支持。

山下真一（Shinichi Yamashita）是 NTT 数据公司 OSS 专业服务部门在日本的首席工程师。他拥有超过 7 年在软件和中间件（Apache、Tomcat、PostgreSQL、Hadoop 的生态系统）工程开发方面的经验。

他在日本也已经出版过了几本关于 Hadoop 的著作。

我要感谢我的同事们。

前言

作为一个开源、分布式的大数据存储系统，HBase 可以处理记录数达几十亿条、字段数达数百万个的大数据，并且只需依靠一些由商品化硬件而构成的集群就可以运行。如果你正在寻找一种能够实时存储和访问海量数据的方法，那么你就必须研究一下 HBase。

《HBase 管理指南》一书提供了一些实用的例子和一些简单详尽的操作步骤，使你可以轻松地管理 HBase。本书的章节涵盖了在云计算环境中管理一个完全分布式的高可用 HBase 集群所需的各种过程。处理如此海量的数据意味着建立起一套有组织、可管理的过程至关重要，本书将帮助你做到这一点。

本书是一部实用性的操作指南，它首先会讲解如何建立一个完全分布式的 HBase 集群以及如何将数据转移到该集群中。你还将学到如何使用各种工具来完成日常管理工作，以及有效地管理和监控集群，以使集群达到最佳的性能。理解 Hadoop 与 HBase 之间的关系能帮你最好地利用 HBase，因此本书也将讲解如何安装 Hadoop 集群，如何配置 Hadoop 使之可与 HBase 进行协作、以及如何调整它们的性能。

1. 本书内容

第 1 章 HBase 集群安装：本章将介绍 HBase 集群的安装方法，首先将介绍基本的单机运行模式的 HBase 实例的安装方法，然后介绍如何在 Amazon EC2 上安装完全分布式运行的高可靠性 HBase 集群。

第 2 章 数据迁移：本章将从简单的任务开始，首先介绍如何使用 Put AP 将数据从 MySQL 导入到 HBase 中；然后，介绍如何使用 importtsv 和批量加载工具将 TSV 数据文件

加载到 HBase 中。我们还将使一个 MapReduce 的例子来导入其他文件格式的数据。其中包括将数据直接放入到一张 HBase 表中，然后在将其写到 HDFS（Hadoop 分布式文件系统）文件系统内的 HFile 格式的文件中。本章的最后一节将介绍如何在将数据加载到 HBase 之前预创建一些区域。

本章附带有几个用 Java 编写的示例代码。我们假设你已具备了一些基本的 Java 知识，所以本章没有解释如何对各节中的 Java 示例代码进行编译和打包。

第 3 章使用管理工具：本章将介绍一些管理工具的用法，比如 HBase 的 Web 用户界面、HBase Shell、HBase hbck 等。我们会解释这些工具的用途以及如何使用它们来完成特定的任务。

第 4 章 HBase 数据备份和恢复：在本章中，我们将介绍如何使用不同方法来对 HBase 的数据进行备份以及各种方法的优点和缺点，还会介绍如何根据你的数据集的规模、资源 and 需求来选择应该使用何种方法。

第 5 章监控及诊断：本章将介绍如何使用 Ganglia、OpenTSDB、Nagios 以及其他一些工具来对 HBase 集群进行监控和故障诊断。我们还是从一个简单的任务开始，首先介绍如何显示 HBase 表的磁盘利用率；然后，我们将安装和配置 Ganglia 来监控 HBase 的各项指标，并展示一个使用 Ganglia 图表的例子。我们还会安装 OpenTSDB，它与 Ganglia 类似但更具可伸缩性，因为它构建在 HBase 之上。

我们将安装 Nagios 来检查一切我们想要检查的指标，包括 HBase 的相关守护进程的健康状况、Hadoop/HBase 的日志、HBase 的不一致问题、HDFS 的健康状况和磁盘空间的利用率。

热点区域问题是一个频繁被人问及的问题，本章的最后一节将介绍一种诊断和修复它的方法。

第 6 章维护和安全：本章前 6 节的内容主要针对 HBase 的各种维护工作，比如发现并纠正故障点、改变集群的规模、进行配置修改等。

本章还将讨论安全性的问题。本章的后三节将介绍安装 Kerberos 并使用 Kerberos 来设置 HDFS 的安全机制，并且最终建立一种安全的 HBase 客户端访问方式。

第 7 章故障排查：本章将研究几个最常遇到的问题。我们将介绍这些问题的错误消息、发生原因以及如何使用故障排除工具来解决这些问题。

第 8 章基本性能调整：本章将介绍如何对 HBase 进行调优，以获得更好的性能。本章还将包括一些对其他调优点进行调优的内容，比如 Hadoop 的配置、JVM 垃圾回收的设置以及操作系统的内核参数等。

第 9 章高级配置和调优：这是本书中又一个关于性能优化的章节。第 8 章介绍了一些通过调整 Hadoop、操作系统设置、Java 和 HBase 本身来改进 HBase 集群的整体性能的方法。这是可适用于很多场合的一般性改进方法。本章将介绍一些更具针对性的调优方法，其中有些适用于写密集的集群，而另一些则以提高集群的读性能为目的。

2. 阅读本书前的准备工作

本书的每一节都会列出你需要的所有东西。

本书所需软件的基本清单如下：

- ◆ Debian 6.0.1 (squeeze)
- ◆ Oracle JDK (Java Development Kit) SE 6
- ◆ HBase 0.92.1
- ◆ Hadoop 1.0.2
- ◆ ZooKeeper 3.4.3

3. 本书读者

本书的读者是 HBase 的管理员和开发人员，同时本书对 Hadoop 的管理员也有所帮助。读者并非一定要有 HBase 的使用经验，但应该对 Hadoop 和 MapReduce 有基本的了解。

4. 排版约定

本书使用了一些不同样式的文本来区分不同种类的信息。下面是这些不同样式的一些例子以及对于它们各自含义的解释。

正文中的代码将以这种方式显示：“可以使用 `stop-hbase.sh` 脚本来关闭 Hbase”。

一段代码将以如下方式显示：

```
nameserver 10.160.49.250 #private IP of ns
search hbase-admin-cookbook.com #domain name
```

当需要你注意一段代码的某一部分时，我们会把相关的行或项以粗体的方式显示出来：

```
MAJOR_COMPACTION_KEY = \x00
```

```
MAX_SEQ_ID_KEY = 96573
TIMERANGE = 1323026325955....1323026325955
hfile.AVG_KEY_LEN = 31
hfile.AVG_VALUE_LEN = 4
hfile.COMPARATOR = org.apache.hadoop.hbase.KeyValue$KeyComparator
```

所有命令行的输入或输出都以如下形式显示：

```
$ bin/ycsb load hbase -P workloads/workloada -p columnfamily=f1 -p
recordcount=1000000 -p threadcount=4 -s | tee -a workloada.dat
YCSB Client 0.1
Command line: -db com.yahoo.ycsb.db.HBaseClient -P workloads/workloada -p
columnfamily=f1 -p recordcount=1000000 -p threadcount=4 -s -load
Loading workload...
```

新术语和重要的文字会以粗体显示。你在电脑屏幕上看到文字（比如在菜单或对话框中出现的文字）也将会以粗体显示，比如：“在 **AWS 管理控制台** 中验证其已经启动”。



警告或重要的说明将会显示在这样一个方框中。



提示和技巧将会显示在这样一个方框中。

5. 读者反馈

我们随时欢迎来自本书读者的反馈。请让我们知道您的想法——您喜欢或可能不喜欢本书的哪些内容。读者的反馈非常重要，它能指引我们撰写出真正对读者有价值的书籍。

对于一般性的反馈，您只需将电子邮件发送至 feedback@packtpub.com 就可以了，请在您的邮件标题中注明本书的名字。

如果您具有某一领域的专长，并且有兴趣写一本书或者对某一本书作出些贡献，那么请您阅读一下我们的作者指南（www.packtpub.com/authors）。

6. 客户支持

我们还为您提供了如下一些东西来帮您从本次购买中获得最大的收益。

6.1 下载示例代码

您可以从您在 <http://www.packtpub.com> 的账户中下载到您所购买的所有 Packt 图书的示例代码文件。如果您是从其他地方购买的本书，那么您可以访问 <http://www.packtpub.com/support> 进行注册，然后我们会以电子邮件的方式直接将这些文件发送给您。

6.2 勘误

虽然我们已尽力确保本书内容的准确性，但错误依旧在所难免。如果你发现了书中的错误——无论是文字还是代码中的错误，都请报告给我们，我们将不胜感激。您这样做可以避免其他读者遭受挫折，并且能帮助我们改进本书的后续版本。如果您发现了错误，请以这种方式报告给我们：访问 <http://www.packtpub.com/support>，选择您所购买的书籍，点击链接“**errata submission form**”，然后输入您勘误的详细内容。在验证完你所提交的勘误之后，我们将接受您的提交并将此勘误上传到我们的网站中或添加到该书勘误区的已有勘误列表中。

6.3 盗版行为

互联网上的盗版问题是所有媒体都在面临的一个将长期存在的问题。Packt 出版公司非常重视保护自己的版权和许可权。如果您在互联网上遇到了本公司作品的任何形式的非法拷贝，都请您立即将其地址或网站名称提供给我们，以便于我们提出赔偿。

请将涉嫌盗版的链接发送至 copyright@packtpub.com 与我们联系。

我们非常感谢您们在保护作者和出版公司方面所提供的帮助，这样我们才能继续为您提供一些更有价值的内容。

6.4 问题

无论您对本书有任何方面的问题，都可以通过 questions@packtpub.com 与我们联系，我们将尽最大努力来解决您的问题。

目录

第 1 章 HBase 集群安装	1
1.1 简介	1
1.2 快速入门	3
1.3 Amazon EC2 的安装及准备	7
1.4 安装 Hadoop	12
1.5 ZooKeeper 安装	17
1.6 修改内核参数设置	20
1.7 HBase 安装	22
1.8 Hadoop/ZooKeeper/HBase 基本配置	26
1.9 安装多个高可用性 (HA) 的主节点	29
第 2 章 数据迁移	42
2.1 简介	42
2.2 通过客户端程序导入 MySQL 数据	43
2.3 使用批量加载工具导入 TSV 文件的数据	49
2.4 编写自定义 MapReduce 任务来导入数据	55
2.5 在数据移入 HBase 前预创建区域	61
第 3 章 使用管理工具	65
3.1 简介	65
3.2 HBase 主 Web 界面	66
3.3 使用 HBase Shell 管理表	69

3.4	使用 HBase Shell 访问 HBase 中的数据	72
3.5	使用 HBase Shell 管理集群	76
3.6	在 HBase Shell 中执行 Java 方法	80
3.7	行计数器	82
3.8	WAL 工具——手动分割和转储 WAL	85
3.9	HFile 工具——以文本方式查看 HFile 的内容	90
3.10	HBase hbck——检查 HBase 集群的一致性	93
3.11	HBase Hive——使用类 SQL 语言查询 HBase 中的数据	96
第 4 章	HBase 数据备份及恢复	104
4.1	简介	104
4.2	使用 distcp 进行关机全备份	105
4.3	使用 CopyTable 在表间复制数据	110
4.4	将 HBase 表导出为 HDFS 上的转储文件	113
4.5	通过从 HDFS 导入转储文件来恢复 HBase 数据	117
4.6	备份 NameNode 元数据	120
4.7	备份区域开始键	124
4.8	集群复制	127
第 5 章	监控与诊断	133
5.1	简介	133
5.2	显示 HBase 表的磁盘利用率	134
5.3	安装 Ganglia 来监控 HBase 集群	136
5.4	OpenTSDB——使用 HBase 监控 HBase 集群	143
5.5	安装 Nagios 来监控 HBase 进程	151
5.6	使用 Nagios 检查 Hadoop/HBase 日志	158
5.7	使用一些简单脚本来报告集群状态	164
5.8	热点区域——诊断写操作	168
第 6 章	维护和安全	173
6.1	简介	173
6.2	启用 HBase RPC 的 DEBUG 级日志功能	174
6.3	平稳节点停机	177

6.4	为集群添加节点	180
6.5	滚动重启	182
6.6	管理 HBase 进程的简单脚本	186
6.7	简化部署的简单脚本	189
6.8	对 Hadoop 和 HBase 进行 Kerberos 身份认证	191
6.9	配置 HDFS 使用 Kerberos 安全保护机制	196
6.10	HBase 的安全保护配置	205
第 7 章	故障排查	210
7.1	简介	210
7.2	故障排查工具介绍	211
7.3	处理 XceiverCount 错误	216
7.4	处理“打开的文件过多”的错误	217
7.5	处理“无法创建新本地线程”错误	219
7.6	处理“HBase 忽略了 HDFS 的客户端配置”问题	221
7.7	处理 ZooKeeper 客户端的连接错误	223
7.8	处理 ZooKeeper 会话过期错误	225
7.9	处理 EC2 上 HBase 的启动错误	228
第 8 章	基本性能调整	237
8.1	简介	237
8.2	设置 Hadoop 分散磁盘 I/O	239
8.3	使用网络拓扑结构脚本使 Hadoop 可感知机架	241
8.4	以 noatime 和 nodiratime 方式装载磁盘	244
8.5	将 vm.swappiness 设为 0 以避免交换	246
8.6	Java GC 和 HBase 堆的设置	248
8.7	使用压缩	252
8.8	管理合并	255
8.9	管理区域分割	257
第 9 章	高级配置和调整	261
9.1	简介	261
9.2	使用 YCSB 对 HBase 集群进行基准测试	262

9.3 增加区域服务器的处理线程数.....	270
9.4 使用自定义算法预创建区域.....	271
9.5 避免写密集集群中的更新阻塞.....	277
9.6 调节 MemStore 内存大小.....	280
9.7 低延迟系统的客户端调节.....	281
9.8 配置列族的块缓存.....	284
9.9 调高读密集集群的块缓存大小.....	287
9.10 客户端扫描类的设置.....	289
9.11 调整块大小来提高寻道性能.....	291
9.12 启用 Bloom 过滤器提高整体吞吐量.....	293