

图书基本信息

书名：《HBase管理指南》

13位ISBN编号：9787115319814

10位ISBN编号：7115319812

出版时间：2013-7-1

出版社：人民邮电出版社

作者：Yifeng Jiang

页数：297

译者：苏南

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.tushu111.com

《HBase管理指南》

内容概要

《HBase管理指南》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。《HBase管理指南》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。

《HBase管理指南》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，他加入了目前这家公司，并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，该团队负责着多个Hadoop/HBase集群的运维工作。

书籍目录

第1章 HBase集群安装

- 1.1 简介
- 1.2 快速入门
- 1.3 Amazon EC2的安装及准备
- 1.4 安装Hadoop
- 1.5 ZooKeeper安装
- 1.6 修改内核参数设置
- 1.7 HBase安装
- 1.8 Hadoop/ZooKeeper/HBase基本配置
- 1.9 安装多个高可用性（HA）的主节点

第2章 数据迁移

- 2.1 简介
- 2.2 通过客户端程序导入MySQL数据
- 2.3 使用批量加载工具导入TSV文件的数据
- 2.4 编写自定义MapReduce任务来导入数据
- 2.5 在数据移入HBase前预创建区域

第3章 使用管理工具

- 3.1 简介
- 3.2 HBase主Web界面
- 3.3 使用HBase Shell管理表
- 3.4 使用HBase Shell访问HBase中的数据
- 3.5 使用HBase Shell管理集群
- 3.6 在HBase Shell中执行Java方法
- 3.7 行计数器
- 3.8 WAL工具——手动分割和转储WAL
- 3.9 HFile工具——以文本方式查看HFile的内容
- 3.10 HBase hbck——检查HBase集群的一致性
- 3.11 HBase Hive——使用类SQL语言查询HBase中的数据

第4章 HBase数据备份及恢复

- 4.1 简介
- 4.2 使用distcp进行关机全备份
- 4.3 使用CopyTable在表间复制数据
- 4.4 将HBase表导出为HDFS上的转储文件
- 4.5 通过从HDFS导入转储文件来恢复HBase数据
- 4.6 备份NameNode元数据
- 4.7 备份区域开始键
- 4.8 集群复制

第5章 监控与诊断

- 5.1 简介
- 5.2 显示HBase表的磁盘利用率
- 5.3 安装Ganglia来监控HBase集群
- 5.4 OpenTSDB——使用HBase监控HBase集群
- 5.5 安装Nagios来监控HBase进程
- 5.6 使用Nagios检查Hadoop/HBase日志
- 5.7 使用一些简单脚本来报告集群状态
- 5.8 热点区域——诊断写操作

第6章 维护和安全

- 6.1 简介
- 6.2 启用HBase RPC的DEBUG级日志功能
- 6.3 平稳节点停机
- 6.4 为集群添加节点
- 6.5 滚动重启
- 6.6 管理HBase进程的简单脚本
- 6.7 简化部署的简单脚本
- 6.8 对Hadoop和HBase进行Kerberos身份认证
- 6.9 配置HDFS使用Kerberos安全保护机制
- 6.10 HBase的安全保护配置
- 第7章 故障排查
 - 7.1 简介
 - 7.2 故障排查工具介绍
 - 7.3 处理XceiverCount错误
 - 7.4 处理“打开的文件过多”的错误
 - 7.5 处理“无法创建新本地线程”错误
 - 7.6 处理“HBase忽略了HDFS的客户端配置”问题
 - 7.7 处理ZooKeeper客户端的连接错误
 - 7.8 处理ZooKeeper会话过期错误
 - 7.9 处理EC2上HBase的启动错误
- 第8章 基本性能调整
 - 8.1 简介
 - 8.2 设置Hadoop分散磁盘I/O
 - 8.3 使用网络拓扑结构脚本使Hadoop可感知机架
 - 8.4 以noatime和nodiratime方式装载磁盘
 - 8.5 将vm.swappiness设为0以避免交换
 - 8.6 Java GC和HBase堆的设置
 - 8.7 使用压缩
 - 8.8 管理合并
 - 8.9 管理区域分割
- 第9章 高级配置和调整
 - 9.1 简介
 - 9.2 使用YCSB对HBase集群进行基准测试
 - 9.3 增加区域服务器的处理线程数
 - 9.4 使用自定义算法预创建区域
 - 9.5 避免写密集集群中的更新阻塞
 - 9.6 调节MemStore内存大小
 - 9.7 低延迟系统的客户端调节
 - 9.8 配置列族的块缓存
 - 9.9 调高读密集集群的块缓存大小
 - 9.10 客户端扫描类的设置
 - 9.11 调整块大小来提高寻道性能
 - 9.12 启用Bloom过滤器提高整体吞吐量
 -

编辑推荐

详尽的操作步骤，真实的管理案例，帮助读者轻松管理HBase。安装一个完全分布式的高可用HBase集群。使用常规的客户端API/或自定义的mapreduce作业来将数据加载到集群中。

精彩短评

- 1、很详细，适合弄个虚拟机慢慢跟着操作。
- 2、3.5分，还是不错的，虽然有点罗嗦，但很多地方还是讲到了，都是经验之谈，如果能再深入就好了
- 3、这本书拿到手之前有点欣喜，但是后来发现讲得东西比较浅显，基本上就是概念-应用，但是缺少实际场景。感觉作用不太大，不如hbase权威指南。。。
- 4、一般 感觉里面没说太多特殊的東西 都是一些常规东西
- 5、内容还挺全的，翻翻留个印象吧
- 6、书籍偏向于运维多一些,对于应用场景的分析几乎没有

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.tushu111.com