

# HADOOP HACKS™

专家使用的实践技巧



O'REILLY®  
中国电力出版社

中野猛 山下真一 猿田浩輔  
上新卓也 小林隆 著  
常娜 陈新 译

---

# Hadoop Hacks(中文版)

中野猛 山下真一 猿田浩輔 上新卓也 小林隆 著  
常娜 陈新 译

**O'REILLY®**

*Beijing • Cambridge • Farnham • Köln • Sebastopol • Tokyo*

O'Reilly Media, Inc. 授权中国电力出版社出版

中国电力出版社

## 图书在版编目 (CIP) 数据

Hadoop Hacks/ (日) 中野猛等著; 常娜、陈新译. —北京: 中国电力出版社, 2014.1

书名原文: Hadoop Hacks

ISBN 978-7-5123-4635-2

I. ①H… II. ①中… ②常… ③陈… III. ①数据处理软件 IV. ①TP274  
中国版本图书馆CIP数据核字 (2013) 第143534号

北京市版权局著作权合同登记

图字: 01-2012-7972号

©2012 by O'Reilly Japan, Inc.

Simplified Chinese Edition, jointly published by O'Reilly Japan, Inc. and China Electric Power Press, 2013. Authorized translation of the Japanese edition, 2012 O'Reilly Japan, Inc., This translation is published and sold by permission of O'Reilly Japan, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

日文原版由O'Reilly Japan Inc. 出版2012。

简体中文版由中国电力出版社出版 2013。日文原版的翻译得到O'Reilly Japan, Inc.的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——O'Reilly Japan, Inc.的许可。

版权所有, 未得书面许可, 本书的任何部分和全部不得以任何形式复制。

书 名/ Hadoop Hacks (中文版)

书 号/ ISBN 978-7-5123-4635-2

责任编辑/ 刘焱

出版发行/ 中国电力出版社 (www.cepp.sgcc.com.cn)

地 址/ 北京市东城区北京站西街19号 (邮政编码100005)

经 销/ 全国新华书店

印 刷/ 汇鑫印务有限公司

开 本/ 787毫米×1092毫米 16开本 24.625印张 455千字

版 次/ 2014年1月第一版 2014年1月第一次印刷

印 数/ 0001-3000册

定 价/ 58.00元 (册)

# O'Reilly Media, Inc.介绍

O'Reilly Media通过图书、杂志、在线服务、调查研究和会议等方式传播创新知识。自1978年开始，O'Reilly一直都是前沿发展的见证者和推动者。超级极客们正在开创着未来，而我们关注真正重要的技术趋势——通过放大那些“细微的信号”来刺激社会对新科技的应用。作为技术社区中活跃的参与者，O'Reilly的发展充满了对创新的倡导、创造和发扬光大。

O'Reilly为软件开发人员带来革命性的“动物书”；创建第一个商业网站（GNN）；组织了影响深远的开放源代码峰会，以至于开源软件运动以此命名；创立了Make杂志，从而成为DIY革命的主要先锋；公司一如既往地通过多种形式缔结信息与人的纽带。O'Reilly的会议和峰会集聚了众多超级极客和高瞻远瞩的商业领袖，共同描绘出开创创新产业的革命性思想。作为技术人士获取信息的选择，O'Reilly现在还将先锋专家的知识传递给普通的计算机用户。无论是通过书籍出版，在线服务或者面授课程，每一项O'Reilly的产品都反映了公司不可动摇的理念——信息是激发创新的力量。

## 业界评论

“O'Reilly Radar博客有口皆碑。”

——Wired

“O'Reilly凭借一系列（真希望当初我也想到了）非凡想法建立了数百万美元的业务。”

——Business 2.0

“O'Reilly Conference是聚集关键思想领袖的绝对典范。”

——CRN

“一本O'Reilly的书就代表一个有用、有前途、需要学习的主题。”

——Irish Times

“Tim是位特立独行的商人，他不光放眼于最长远、最广阔的视野并且切实地按照Yogi Berra的建议去做了：‘如果你在路上遇到岔路口，走小路（岔路）。’回顾过去Tim似乎每一次都选择了小路，而且有几次都是一闪即逝的机会，尽管大路也不错。”

——Linux Journal

---

# 致谢

## 作者简介

**中野 猛 (nakano takeshi)**

1976年出生于兵库县。从奈良先端大学毕业后进入招聘股份公司工作。在信息系统部门MIT中，从事Web基础设施的构建以及R25等网站建立、开发、运用等。期间，推进了Solr等OSS的导入和高速缓存、中间件开发等。近两年正在以Hadoop为中心进行探索研究。爱好是潜水（最喜欢的地方是墨西哥、拉巴斯）。twitter: @tf0054

**山下 真一 (yamashita sinichi)**

NTT股份公司数据专员。进入公司以来一直从事基于OSS的基础系统的配置、构建、评价。最近增加了以Hadoop为中心的基础系统的配置、构建、评价的研究，正在研究效率更高的处理MapReduce的方法。相比Hive来说，更偏好Pig。

**猿田 浩辅 (saruta kousuke)**

NTT股份公司数据勤务。2009年进入公司以来，从事以Hadoop为开端的OSS的导入支持和技术验证与开发工作。目前，不仅构建了规模达数百台的Hadoop集群，还在实施应用程序的设计支持等。最近进行从RDBMS向Hive的业务移动和Pig的导入支持等工作。特别喜欢猫，爱吃甜食，自称“甜点工程师”。

**上新 卓也 (ueshin takuya)**

从事过网络开发公司中的广告系统和存取解析，后来成为自由职业

者。通过工作了解到Hadoop，特别热衷于HBase，在自己家中正在使用HBase集群。目前就职于Nautilus技术股份公司。twitter: @ueshin

#### 小林 隆 (kobayeshi ryu)

最近主要从事使用cloud的开发、使用Hadoop的MapReduce的开发，以及使用NoSQL的开发。另外，还开发了Cassandra开源代码的GUI (<http://code.google.com/a/apache-extras.org/p/cassandra-gui/>)。翻译作品有“Cassandra” (O'Reilly) 等。twitter: @ryu\_kobayashi

## 关于合作者

#### 向井 育正 (mukai ikumasa)

从事业务系统、嵌入式系统、Web系统等系统的开发。不拘泥于小节，自称“职业软件开发人员”。他的座右铭是，“即使乌龟走路很慢，但像乌龟那样一步一个脚印踏踏实实走下去，也会不断前进。”一直虚心学习。twitter: @kingdom9999

#### 饭田 慎司 (ida shinji)

最初是应用程序开发工程师，现在擅长中间件开发。也擅长Hadoop，喜欢KVM。现在又喜欢上了ZooKeeper。twitter: @siniida

#### 大坪 正典 (otsubo masanori)

新日铁解决方案股份公司研究员。在大阪大学专心研究数据挖掘。完成学业后，积累了HPC的经验，对分布系统产生兴趣，2009年进入Hadoop的世界。2011年获取CCDH。爱好是飞镖和围棋。每次看到米奇商标的LOGO总是会看成棋子，真没办法。twitter: @otsubo0423

#### 小林 孝 (kobayashi takashi)

从项目的经营到编程，小林孝在所有领域都运用自如，在业务上是高级工程师，在家庭上是一个好爸爸。在业界经历早就超过了20年，打破了IT退休年龄的限制。目前正在从事以利用开源代码软件开发为中心的活动。twitter: @tinyforest\_1969

#### 安斋 佑司 (anzai yuji)

擅长Web应用程序开发，希望将数据解析结果显示到图形的工程师。喜欢可以免费使用的程序库。认为表示方法和内容一样重要。

#### 石川 信行 (ishikawa nobuyuki)

2009年进入招聘股份公司。开发过基于Java的营业系统、使用Javascript的智能手机网站，目前以招聘的很多企业为对象从事基于Hadoop的数

据分析改进业务。爱好是饲养独角仙、海水鱼，潜水等亲近自然的  
活动。twitter: @ground\_beetle

#### 石浜 智也 (ishihama tomoya)

Light Cafe股份公司技术负责人。喜欢Ruby、Vim、Mac和电  
影。平常从事Hadoop项目开发。休息日学习Titanium Mobile。  
twitter: @ishmt

#### 太田 一树 (ota kazuki)

东京大学信息理工学系研究科计算机科学硕士毕业。目前，在美  
国Treasure Data, Inc.担任CTO。对Data Intensive Computing/分布  
式并行技术非常感兴趣。

#### 金卷 吉伸 (kanemaki yoshinobu)

自2010年起在Light Cafe股份公司担任Java首席工程师。最近，正  
在犹豫是否要参与Ruby。另外，正燃烧着抚育幼儿的热情。

#### 箭内 直树 (yanai naoki)

正在进行网络广告系统的产品开发。使用Hadoop，处理每天积  
累的大量广告发送日志。最近经常参加数据挖掘方面的社区。  
twitter: @yanaoki

#### 田井中 智也 (tainaka tomoya)

NTT股份公司数据服务工程师。2010年进入公司以来，从事  
Hadoop的构建、运用服务，经历了数百台规模的Hadoop集群构  
建。爱好是旅行或露营等户外运动，最近躲在家里接触PC的时候  
多了，感到有危机感。

#### 嶋内 翔 (shimauchi syo)

Cloudera股份公司客户运营工程师。2011年4月作为第一个日本  
工程师参加Cloudera。作为支持代表，主要是对日本的顾客提  
供服务。作为工程师在NEC的开源代码团队中工作5年，参与了  
Linux系统的设计、构建等业务。最近利用业余时间，向Hadoop  
community投稿补丁。工作上使用Java，但喜欢的语言是Python。  
爱好是在twitter中为“hadoop”或“hbase”等关键词感到困惑的  
人给出种种建议。twitter: @shiumachi

### **草薙 昭彦 (kushanagi akihiko)**

在以前的Sun Microsystems公司中，作为Java传播者从事Java技术的普及活动或系统设计、导入支持业务。以前作为游戏程序员也很活跃。目前在EMC中，致力于Greenplum产品的设计以及推广活动。喜欢豆芽菜。twitter: @nagix

### **河村圭介 (kawamura keisuke)**

出生于北海道钏路市的软件工程师。在乐天股份公司工作时接触到Hadoop。目前在Glee股份公司从事社交网络服务的Good数据分析系统的开发，继续关注Hadoop。主要负责本书的策划、跟进与协调。twitter: @kkawamura

# 目录

|                            |    |
|----------------------------|----|
| 前言 .....                   | 1  |
| 第1章 系统架构/运用技巧 .....        | 5  |
| 运行HDFS环境的参数 .....          | 6  |
| 运行MapReduce环境需要的参数 .....   | 7  |
| 总结 .....                   | 8  |
| 文件描述符的设置 .....             | 9  |
| Java的安装 .....              | 10 |
| 总结 .....                   | 11 |
| 本技巧中介绍的HA的构成 .....         | 11 |
| HA集群的构建过程 .....            | 13 |
| 疑难解答 .....                 | 31 |
| 总结 .....                   | 33 |
| 可以获取的统计信息 .....            | 34 |
| 总结 .....                   | 40 |
| 关于CDH3同一版本间的更新 .....       | 41 |
| 总结 .....                   | 42 |
| 准备 .....                   | 43 |
| 理解操作 .....                 | 43 |
| 使用Oracle的操作确认 .....        | 45 |
| 总结 .....                   | 52 |
| Sqoop的PostgreSQL联合功能 ..... | 53 |
| 在PostgreSQL中的使用 .....      | 53 |
| PostgreSQL联合的挑战 .....      | 55 |

|                                    |           |
|------------------------------------|-----------|
| 总结 .....                           | 55        |
| 什么是Azkaban .....                   | 55        |
| Azkaban的安装 .....                   | 56        |
| 总结 .....                           | 61        |
| 作业的定制 .....                        | 61        |
| 总结 .....                           | 66        |
| <b>第2章 应用程序开发技巧 .....</b>          | <b>67</b> |
| 将Hadoop配置文件拷贝到外部应用程序 .....         | 67        |
| 从应用程序操作MapReduce .....             | 68        |
| 总结 .....                           | 69        |
| 什么是InMapperCombiner .....          | 70        |
| InMapperCombiner的实现方法 .....        | 70        |
| 总结 .....                           | 72        |
| 自定义Writable类型主要分为两种 .....          | 73        |
| 如何生成一个Writable接口的子类 .....          | 73        |
| 总结 .....                           | 75        |
| 自定义Partitioner的实现方法 .....          | 75        |
| 总结 .....                           | 77        |
| 什么是DistributedCache .....          | 78        |
| DistributedCache的使用 .....          | 78        |
| 总结 .....                           | 79        |
| 自定义CombineFileInputFormat的实现 ..... | 80        |
| 自定义CombineFileInputFormat的使用 ..... | 84        |
| 总结 .....                           | 86        |
| 获取MRUnit Jar文件 .....               | 87        |
| Map的测试 .....                       | 87        |
| Reduce的测试 .....                    | 88        |
| 测试MapReduce作业 .....                | 88        |
| 总结 .....                           | 89        |
| 生成自定义Writable .....                | 90        |
| 生成GroupingComparatorClass .....    | 90        |
| 生成PartitionerClass .....           | 91        |
| 生成SortComparatorClass .....        | 91        |

|                             |     |
|-----------------------------|-----|
| 设置各个Comparator .....        | 92  |
| 总结 .....                    | 92  |
| 数据的内容 .....                 | 93  |
| Map端连接的代码 .....             | 93  |
| 输出结果 .....                  | 94  |
| 单独读入文件进行连接 .....            | 94  |
| 总结 .....                    | 95  |
| 数据的内容 .....                 | 95  |
| Map的定义 .....                | 95  |
| Reduce的定义 .....             | 96  |
| 作业的定义 .....                 | 96  |
| 输出结果 .....                  | 97  |
| 总结 .....                    | 97  |
| 定义作业 .....                  | 98  |
| 总结 .....                    | 100 |
| InputFormat类 .....          | 100 |
| OutputFormat类 .....         | 102 |
| 总结 .....                    | 104 |
| Cassandra的hadoop程序包内容 ..... | 105 |
| Map类 .....                  | 105 |
| Reduce类 .....               | 106 |
| 定义作业 .....                  | 106 |
| 总结 .....                    | 107 |
| 获取文件系统信息 .....              | 108 |
| 操作HDFS的API .....            | 109 |
| 总结 .....                    | 111 |
| 分析的要点 .....                 | 112 |
| 用计数器可以确认的项目 .....           | 112 |
| 处理结束后的确认方法 .....            | 113 |
| 经过一定时间后的确认方法 .....          | 114 |
| 作业历史的管理方法 .....             | 115 |
| 总结 .....                    | 115 |
| Hadoop中可以处理的压缩格式 .....      | 116 |
| Hadoop中处理lzo压缩的准备 .....     | 116 |

|                                  |            |
|----------------------------------|------------|
| 在MapReduce程序中处理压缩文件 .....        | 117        |
| 在Hive中处理压缩文件 .....               | 120        |
| 总结 .....                         | 121        |
| <b>第3章 HBase技巧 .....</b>         | <b>122</b> |
| importtsv 工具 .....               | 122        |
| completebulkload工具 .....         | 123        |
| 总结 .....                         | 123        |
| sqoop import .....               | 124        |
| 自定义PutTransformer .....          | 124        |
| 总结 .....                         | 125        |
| 用于HFileOutputFormat的Mapper ..... | 126        |
| 作业的设置 .....                      | 126        |
| 总结 .....                         | 127        |
| Java API .....                   | 128        |
| HBase shell .....                | 129        |
| 总结 .....                         | 129        |
| 什么是协处理器 .....                    | 130        |
| Observer .....                   | 130        |
| Endpoint .....                   | 132        |
| 总结 .....                         | 135        |
| 创建自定义Filter .....                | 135        |
| 配置自定义Filter .....                | 137        |
| 总结 .....                         | 138        |
| export工具 .....                   | 138        |
| import工具 .....                   | 139        |
| 总结 .....                         | 139        |
| 复制机制 .....                       | 139        |
| 配置 .....                         | 140        |
| 数据的确认 .....                      | 141        |
| 总结 .....                         | 141        |
| <b>第4章 Hive技巧 .....</b>          | <b>142</b> |
| SQL语句和HiveQL语句的书写方法的不同 .....     | 143        |

|                                      |     |
|--------------------------------------|-----|
| SELECT句中不存在的列不能排序 .....              | 147 |
| NULL的处理 .....                        | 147 |
| ROWNUM关联 .....                       | 148 |
| 总结 .....                             | 148 |
| 标点字符、带圈字符的存在 .....                   | 149 |
| PK的存在 .....                          | 149 |
| 加载数据 .....                           | 149 |
| Join时的数据保持方法 .....                   | 150 |
| WHERE句中的比较条件 .....                   | 150 |
| 置换处理和UDF .....                       | 150 |
| key数和Out Of Memory .....             | 150 |
| 总结 .....                             | 151 |
| 处理单位 .....                           | 151 |
| ORDER BY语句和SORT BY语句 .....           | 151 |
| 输出中间连接表 .....                        | 152 |
| MapReduce任务不发生的处理 .....              | 153 |
| 一次连接多个表 .....                        | 153 |
| 总结 .....                             | 153 |
| 查询高速化的探讨 .....                       | 154 |
| 总结 .....                             | 158 |
| Hive的用户定义函数 .....                    | 159 |
| UDF的安装 .....                         | 159 |
| UDAF的实现 .....                        | 161 |
| 使用用户定义函数 .....                       | 164 |
| 总结 .....                             | 165 |
| 什么是SerDe .....                       | 165 |
| 文本数据的情况（使用RegexSerDe） .....          | 166 |
| 文本数据的情况（使用CSV用的非标准SerDe） .....       | 167 |
| 结构化/二进制数据的场合（使用JSON用的非标准SerDe） ..... | 169 |
| 总结 .....                             | 171 |
| 使用TRANSFORM句 .....                   | 171 |
| 使用UDAF .....                         | 174 |
| 总结 .....                             | 176 |

|                           |     |
|---------------------------|-----|
| 使用注释 .....                | 176 |
| 使用模式 .....                | 177 |
| 使用锁定机制 .....              | 178 |
| 总结 .....                  | 181 |
| 流程 .....                  | 182 |
| 使用方法 .....                | 182 |
| 总结 .....                  | 186 |
| Hive的日志的种类 .....          | 187 |
| 通过Hive会话日志可以确认的内容 .....   | 187 |
| 总结 .....                  | 190 |
| Hive-HDFS-RDBMS的关系 .....  | 190 |
| 向Hive的直接导入 .....          | 191 |
| 文件形式和直接导入的关系 .....        | 193 |
| 列指定·WHERE条件指定 .....       | 195 |
| 指定了Hive的分区key的导入 .....    | 197 |
| 总结 .....                  | 198 |
| Hive→Oracle的差异数据的交换 ..... | 200 |
| 总结 .....                  | 203 |

## 第5章 Pig技巧 ..... 204

|                              |     |
|------------------------------|-----|
| 进行LOAD处理要实现的项目 .....         | 204 |
| 进行STORE处理要实现的项目 .....        | 206 |
| 在一个UDF中实现LOAD和STORE的处理 ..... | 208 |
| 总结 .....                     | 209 |
| 数据的保存方法 .....                | 209 |
| MapReduce作业的评价方法 .....       | 210 |
| 用户单独的处理 .....                | 210 |
| 输出的日志 .....                  | 211 |
| 总结 .....                     | 211 |
| Pig的用户定义函数 .....             | 211 |
| 用户定义函数中可以使用的数据类型 .....       | 212 |
| 用户定义函数的实现和使用方法 .....         | 212 |
| 总结 .....                     | 218 |
| Java中的PigLatin嵌入方法 .....     | 218 |

|                                |            |
|--------------------------------|------------|
| 执行范例 .....                     | 219        |
| 总结 .....                       | 220        |
| DESCRIBE：确认数据结构 .....          | 221        |
| EXPLAIN：确认MapReduce的执行计划 ..... | 222        |
| DUMP：显示处理内容 .....              | 225        |
| 总结 .....                       | 226        |
| 数学函数 .....                     | 226        |
| 字符串操作函数 .....                  | 227        |
| 总结 .....                       | 228        |
| <b>第6章 Mahout技巧 .....</b>      | <b>229</b> |
| 可以做什么 .....                    | 230        |
| 推荐（Recommendation） .....       | 230        |
| 聚类（Clustering） .....           | 230        |
| 分类（Classification） .....       | 231        |
| 其他 .....                       | 231        |
| 总结 .....                       | 231        |
| 安装 .....                       | 232        |
| 执行方法 .....                     | 232        |
| 总结 .....                       | 234        |
| 总结 .....                       | 237        |
| 输入数据的形式 .....                  | 237        |
| 处理的执行 .....                    | 237        |
| 执行命令的内容 .....                  | 238        |
| 与文本数据的对应 .....                 | 239        |
| 总结 .....                       | 247        |
| 输入数据 .....                     | 248        |
| 执行关联分析 .....                   | 248        |
| 执行结果文件 .....                   | 250        |
| 总结 .....                       | 251        |
| 从学习（预测模型）到预测的过程 .....          | 252        |
| 预测模型的构建和应用 .....               | 253        |
| 总结 .....                       | 264        |
| 分布式并行化的必要性 .....               | 264        |

|                              |            |
|------------------------------|------------|
| 随机森林的分布式并行图像 .....           | 264        |
| In-memory模式 .....            | 269        |
| 总结 .....                     | 273        |
| 应用例子 .....                   | 274        |
| 创建方法 .....                   | 274        |
| 创建驱动程序 .....                 | 276        |
| 使用驱动程序 .....                 | 276        |
| 总结 .....                     | 278        |
| K平均法 .....                   | 278        |
| 输入数据的形式 .....                | 278        |
| 输入数据的变换 .....                | 278        |
| 执行kmeans .....               | 279        |
| 获取执行结果 .....                 | 279        |
| 结果文件的内容 .....                | 280        |
| 总结 .....                     | 280        |
| <b>第7章 ZooKeeper技巧 .....</b> | <b>281</b> |
| 客户端程序库 .....                 | 282        |
| 简单的API .....                 | 282        |
| 与ZooKeeper的连接和断开 .....       | 282        |
| znode的创建 .....               | 285        |
| 确认znode的存在 .....             | 288        |
| 从znode读取数据 .....             | 289        |
| znode中写入数据 .....             | 290        |
| 删除znode .....                | 291        |
| 获取子节点的列表 .....               | 291        |
| 总结 .....                     | 293        |
| 认证方案 .....                   | 293        |
| ACL和权限 .....                 | 293        |
| ACL的设置 (setACL) .....        | 294        |
| 获取ACL (getACL) .....         | 296        |
| 总结 .....                     | 297        |
| 操作 .....                     | 298        |
| 总结 .....                     | 307        |

|                           |     |
|---------------------------|-----|
| 启动 .....                  | 307 |
| 连接 .....                  | 308 |
| 操作数据 .....                | 308 |
| 尝试监视 .....                | 310 |
| 完成 .....                  | 311 |
| 总结 .....                  | 312 |
| 体系结构 .....                | 312 |
| Rest服务器的启动 .....          | 312 |
| 总结 .....                  | 316 |
| C客户端程序库 .....             | 319 |
| 总结 .....                  | 320 |
| 什么是tickTime? .....        | 321 |
| 设置方法 .....                | 321 |
| 设置其他tickTime影响的项目 .....   | 321 |
| 以其他tickTime为基准进行的处理 ..... | 324 |
| 总结 .....                  | 325 |
| 什么是ZAB .....              | 326 |
| 使用ZAB的ZooKeeper内部处理 ..... | 326 |
| 两阶段提交 .....               | 327 |
| 总结 .....                  | 329 |
| 构成ensemble .....          | 329 |
| leader选出运算法则 .....        | 332 |
| 总结 .....                  | 333 |
| 状态 .....                  | 334 |
| 总结 .....                  | 338 |
| 分组 .....                  | 339 |
| 权重 .....                  | 340 |
| 总结 .....                  | 342 |
| 分布并行队列的实现 .....           | 342 |
| 总结 .....                  | 351 |
| 什么是BookKeeper .....       | 351 |
| BookKeeper的基本操作 .....     | 354 |
| 尝试BookKeeper .....        | 354 |
| 有效利用BookKeeper .....      | 360 |