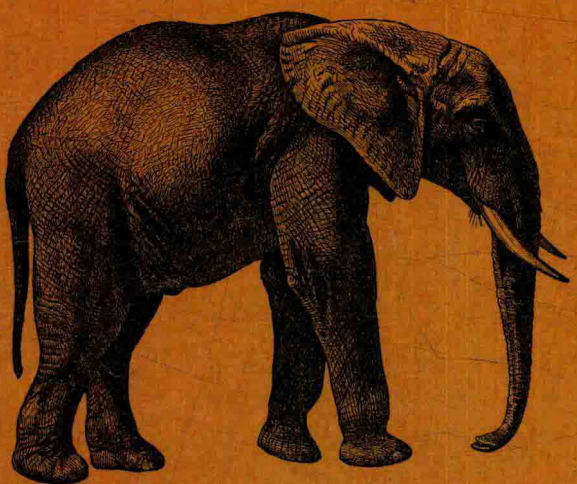
O'REILLY®

Hadoop

生态系统

Field Guide to Hadoop



KEVIN SITTO & MARSHALL PRESSER 著

陈新 唐晓 译

版社

Hadoop生态系统

Kevin Sitto & Marshall Presser 著

陈新 唐晓 译

Beijing • Cambridge • Farnham • Köln • Sebastopol • Tokyo O'REILLY®

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生态系统/（美）凯文·斯托（Kevin Sitto），（美）马歇尔·普瑞斯（Marshall Presser）著；陈新，唐晓译. —北京：中国电力出版社，2016.11

书名原文：Field Guide to Hadoop

ISBN 978-7-5123-9598-5

I. ①H… II. ①凯… ②马… ③陈… ④唐… III. ①数据处理软件 IV. ①TP274

中国版本图书馆CIP数据核字（2016）第174896号

北京市版权局著作权合同登记

图字：01-2016-4203号

封面设计/ Ellie Volckhausen，张健

出版发行/ 中国电力出版社（http://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/ 北京天宇星印刷厂

开　　本/ 787毫米×980毫米　16开本　7.75印张　123千字

版　　次/ 2016年11月第一版　2016年11月第一次印刷

印　　数/ 0001－3000册

定　　价/ 28.00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

感谢我漂亮的妻子Erin，感谢她无限的耐心，也要送给我优秀的孩子们Dominic和Ivy，他们把我照顾得很好。

——Kevin

感谢我的妻子Nancy Sherman，感谢她在我写作、改写和再次改写的过程中不断的鼓励。同样，要非常感谢可爱的小黄象，若没有它，我们也不会想到写这样一本书。

——Marshall

目录

前言

Hadoop 是什么？为什么需要关注它？本书将帮助你理解 Hadoop 是什么，现在先让我们来回答一下第二个问题。Hadoop 是当前最流行的可以支持存储和大数据分析的独立平台。如果你和你的组织计划进入大数据这个激动人心的世界，那么你将必须决定 Apache Hadoop 是否是一个适合使用的平台，同时还需要决定 Hadoop 中哪些组件最适合你的任务。本书的作用就是为你介绍 Hadoop 的各个主题，开启你的 Hadoop 旅程。

现在已经有许多关于 Hadoop 及其相关技术的书籍、网站和培训班。本书与众不同的地方是并没有提供一个关于 Hadoop 特定方面或 Hadoop 生态系统中关于各种组件的冗长的教程介绍。本书并不是一本详细讨论 Hadoop 中任何一个主题的书籍。相反，与鸟语林中的路标一样， 每一章都聚焦一个拥有共同主题的 Hadoop 生态系统中的一部分。在每一章中，简要介绍了相关的技术和主题；我们解释了该技术与 Hadoop 的关系，并讨论了为什么该技术适用某些需求（以及在某些场景下并不适用）。最后，本书包含了多个小节内容，这些小节讨论了多个 Apache Hadoop 的项目和子项目及其相关技术，并给出了相关技术和方法的教程链接。

在每一小节，通常包含类似下面的一个表格：

许可证	< 许可证信息 >
活跃度	< 无，低，中，高 >
用途	< 相关用途 >

官方网站　　< 网址 >

Hadoop 集成度　　完全集成、API 兼容、未集成、不适配

现在详细看看这些类别的描述：

许可证（License）

尽管本书中各个小节中的软件都是开源的，但这些软件使用了不同的许可证，大体上是一样的，但也有一些小的不同。如果你有计划在产品中使用这些软件，那么你就需要对这些许可证的使用条件比较熟悉。

活跃度（Activity）

我们尽可能地度量了该技术在开发中使用的活跃度。可能在某些场景下我们的度量是有错误的，并且活跃度等级可能在撰写完本主题后也会发生改变。

用途（Purpose）

该技术可以用来做什么？我们将尽力根据用途的相似度将主题进行分组，但是有时候我们发现某一个主题可以适用不同章节。生活就是不断做出选择，这就是我们的选择。

官网

对于本书涉及的技术，互联网上最权威的站点可能就是它的官方网站了。

Hadoop 集成度

在写作时，我们并不确定有哪些主题会包含在本书的第一个版本中。最原始清单中的一些主题是紧密集成或者说绑定在 Hadoop 中的。有一些其他技术是可选择的，或者是可以和 Hadoop 共同工作的，但并不是 Apache Hadoop 家庭中的一部分。对于那些情形，我们尽可能描述在写作的时候该技术的集成度。当然随着时间变化，毫无疑问这将会发生改变。

本书并不需要从头到尾阅读。如果你完全是一个新的 Hadoop 用户，那么你可以从第 1 章（介绍）开始阅读。然后可以看看感兴趣的主题，阅读关于该组件的小节内容，阅读章节的章下文，也可以浏览同一章中的其他小节。本书通常包含了指向其他可能相关小节的链接。你也可以查看关于该主题的教程或者官方网站。

我们按照图 P-1 所示的模式，将所有主题分类为多个章节。许多主题纳入 Hadoop Common（之前称为 Hadoop Core），这是支持所有其他 Apache Hadoop 模块的最基础的工具和技术。然而，这些工具在大数据生态系统中也起到了很重要的作用，并不仅仅局限于是 Hadoop 的内核技术。在本书中我们也讨论了许多在大数据视角起了重要作用的相关技术。

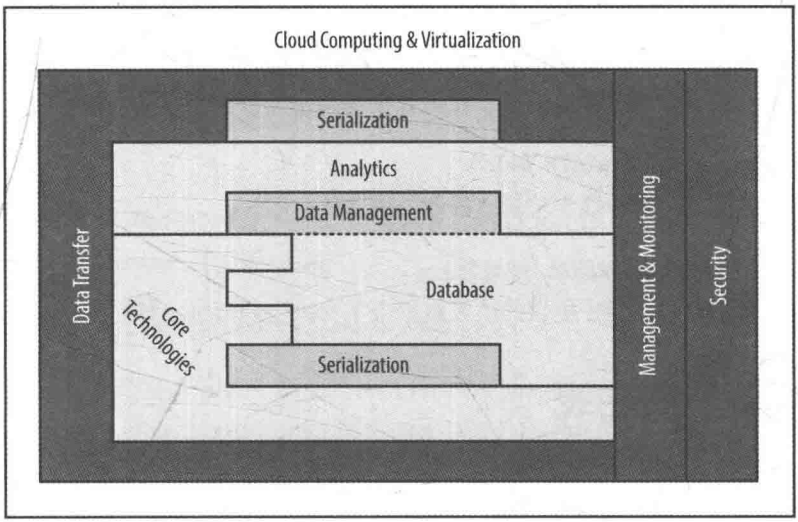


图 P-1：本书涉及主题的概览

在本书中，我们并没有包含任何专用的 Hadoop 发行版本。我们认识到这些项目是重要并且有重大意义的，但是商业系统变化特别快，所以我们重点只关注开源技术。

当前，开源技术在 Hadoop 和大数据市场中有很重要的影响力，许多商业的解决方案都是深度基于本书中描述的开源技术来实现的。如果读者对本书中讨论的开源技术感兴趣，有采用该技术的倾向，也可以去寻找商业发行版本。

本书并不是一个每年或每两年才更新的一个静态文档。我们的目标是尽可能地保持内容最新，随着 Hadoop 环境的增长而增加新的内容，有一些陈旧的技术要么会消失，要么进入维护模式，因为它们可能会被其他能够满足更新技术需求或由于其他原因获得支持的技术所代替。

因为本书涉及的主题可能会变化很快，非常欢迎读者将建议和评论反馈给 Kevin(ksitto@gmail.com) 和 Marshall(bigmaish@gmail.com)。你们有任何建议都可以反馈给我们，我们将十分感谢。

本书使用的排版约定

本书使用如下排版约定：

斜体 (italic)

用来表示新术语、URL、email 地址、文件名、文件扩展名等。

等宽字体（`constant width`）

用来表示程序列表，同时在段落中引用的程序元素（例如变量、函数名、数据库、数据类型、环境变量、声明和关键字等）也用该格式表示。

等宽黑体（`constant width bold`）

用于表示需要用户逐字符输入的命令或其他文本。

等宽斜体（*`constant width italic`*）

用于表示应该以用户提供的值或根据上下文决定的值加以替换的文本。

Safari® 图书在线

Safari® Book Online 是一个按需定制的数字化图书馆，它提供了来自全球技术和业务上最顶尖的作者的专业内容，既包括书籍，也包括视频。

技术专业人员、软件开发者、网页设计者和商业创新专业人员都可以使用 Safari Book Online 来作为研究问题、解决、学习和认证培训的主要资源。

Safari Book Online 为组织者、政府机构和个人提供了各种价格范围的资源。订阅者可以通过一个统一的可搜索数据库访问成千上万的书籍、培训视频和预先出版的手稿，提供资源的出版社包括 O'Reilly Media, Prentice Hall Professional, Addison-Wesley Professional, Microsoft Press, Sams, Que, Peachpit Press, Focal Press, Cisco Press, John Wiley & Sons, Syngress, Morgan Kaufmann, IBM Redbooks, Packt, Adobe Press, FT Press, Apress, Manning, New Riders, McGraw-Hill, Jones & Bartlett, Course Technology 等。如需更多关于 Safari Book Online 的信息，请访问我们的网站。

如何联系我们

美国：

O'Reilly Media，Inc.
1005 Gravenstein Highway North
Sebastopol，CA 95472

中国：

北京市西城区西直门南大街 2 号成铭大厦 C 座 807 室（100035）
奥莱利技术咨询（北京）有限公司

我们为本书提供了网页，该网页上面列出了勘误表、范例和任何其他附加的信息。您可以访问如下网页获得：*http:// bit.ly/field-guide-hadoop*。

要询问技术问题或对本书提出建议，请发送电子邮件至：*bookquestions@oreilly.com*

要获得更多关于我们的书籍、会议、资源中心和 O'Reilly 网络的信息，请参见我们的网站：*http://www.oreilly.com*

我们的 Facebook 地址：*http://facebook.com/oreilly*

我们的 Twitter 地址：*http://twitter.com/oreillymedia*

我们的 YouTube，地址：*http://www.youtube.com/oreillymedia*

致谢

我们要感谢本书的评审 Harry Dolan、Michael Park、Don Miner 和 Q Ethan McCallum。非常感谢他们为此付出的时间，以及洞察力和耐心。

我们还要非常感谢 O'Reilly 团队给我们的帮助。我们特别希望感谢 Mike Loukides，感谢他在我们刚开始写作时给我们提供的宝贵的帮助，还要感谢 Ann Spencer，感谢她帮助我们理清如何写作本书的思路，要感谢 Shannon Cutt，他们提供的评论使得本书成为可能。另外要特别感谢 Rebecca Demarest 和 Dan Fauxsmith。

此外，还要特别感谢 Paul Green，感谢他在大数据还是一个不为人知的小事物时，教会我们什么是大数据，还要感谢 Don Brancato，感谢他强迫一个程序员来阅读 strunk and white 写的《the elements of style》[译注 1]。

译注 1: 这是一本英语语法书籍。

第 1 章

关键技术

在 2002 年，当时万维网（World Wide Web）相对来说还比较新，大家也还没有使用谷歌来搜索东西，Doug Cutting 和 Mike Cafarella 想要抓取网页并对其内容做索引，以此来开发一个因特网的搜索引擎。他们启动了一个名为 Nutch 的项目来做这件事情，该项目需要通过一个可伸缩的方法来存储这些索引的内容。在 2002 年，通常采用关系数据库系统（RDBMS）来组织和存储数据的标准方法，该系统用一种名为 SQL 的语言来存取数据。但是几乎所有的 SQL 和关系存储都不适合因特网上搜索引擎的存储和检索。它们很昂贵，不具备可伸缩性，不能接受失败请求，并且可能达不到所需的性能。

在 2003 年和 2004 年，谷歌发表了两篇重要的论文，一篇是有关 Google 文件系统的（*http://bit.ly/1CgWGTy*），另一篇是关于集群服务器上名为 MapReduce（*http://bit.ly/12c3Ifq*）的编程模型的。Cutting 和 Cafarella 将这些技术应用到他们的项目中，最终 Hadoop 就产生了。Hadoop 不是一个首字母的缩写。Cutting 的儿子有一个黄色的毛绒大象，它的名字叫 Hadoop，然后不知怎么，这个名字就被标记在了这个项目上，并且这只伶俐可爱的小象就成为它的图标。Yahoo! 最开始将 Hadoop 作为它搜索引擎的基础，然后很快就扩展到了很多其他的组织。现在 Hadoop 是最主要的大数据平台。已有很多详细描述 Hadoop 的资源。在这里将看到许多有关组件的概要描述，通过它们的指向还可以学习到更多的相关信息。

Hadoop 由三个主要的资源组成：

- Hadoop 分布式文件系统 (HDFS)。
- MapReduce 编程平台。
- Hadoop 生态系统，这是一个工具的集合，它们使用或在 MapReduce 和 HDFS 周围去存储和组织数据，以及对那些运行 Hadoop 的机器进行管理。

这些机器被称为一个集群，也就是一个服务器组，它们几乎总是运行在一些不同版本的 Linux 操作系统上去执行共同的任务。

Hadoop 生态系统由模块组成，这些模块帮助编写系统、管理及配置集群、管理集群中的数据和存储、执行分析任务，以及类似的工作。本书中的大部分模块将描述生态系统中的不同组件及相关的技术。

1.1 Hadoop 分布式文件系统（HDFS）

许可证	Apache License, Version 2.0
活跃度	高
用途	大容量、容错性、可存储非常大的数据集的廉价存储
官方网站	http://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-hdfs/HdfsUserGuide.html
Hadoop 集成度	完全集成

Hadoop 分布式文件系统（HDFS）是 Hadoop 集群中的一部分，它是用来存储数据的。为了建造一个数据密集型应用，HDFS 的目的是在廉价的商业服务器上运行集群。HDFS 对高性能、阅读密集型操作进行了优化，且对集群上的错误有适应性。它不能阻止错误的发生，但是它也不太可能丢失数据，因为 HDFS 在默认情况下对它的每个数据块都存有多个副本。除此之外，HDFS 是一次写入、多次读取的文件系统：一旦一个文件被创建，文件系统 API 仅仅允许你去添加文件，而不允许你去重新改写它。因此，通常 HDFS 对在线事务处理系统（OLTP）不太适用。HDFS 最主要的用途是用于连续读取大型文件。这些文件将分解成大文件块，通常是 64MB 或者更大，这些块分布在服务器的各个节点之中。

HDFS 文件系统不兼容 POSIX，就好像你在 Linux、Mac OS X，以及一些 Windows 平台上看到的那样（可以查看 POSIX 维基百科（*http://bit.ly/16TI2GO*）获取相关的简要说明）。它不由服务器节点上的操作系统内核来管理。HDFS

上的块映射到主机底层文件系统的文件上，通常采用的是 Linux 系统的 ext3。HDFS 假定主机上底层的磁盘没有被 RAID 所保护，所以默认情况下，每个块都有三个备份，并且分布在不同的集群节点上。这就在节点或者磁盘损坏的时候为数据提供了保护，目的是为了防止数据丢失，并且有助于对 Hadoop 上的数据进行访问，而不是将这些数据通过网络移动后再访问它们。

尽管这些说明超过的本书的范围，HDFS 上有关文件的元数据是通过 NameNode 来管理的，Hadoop 相当于 UNIX/Linux 的超级块。

1.1.1 教程链接

通常情况下，你可以通过其他的一些工具与 HDFS 进行交互，如 Hive（详见第 2 章）或者 Pig（详见第 5 章）。即便如此，你想立刻用 HDFS 来工作还需要一些时间；Yahoo！已经出版了一本对这个基础系统进行配置和探索的非常优秀的指导说明（*http://yhoo.it/1uEUNQJ*）。

1.1.2 示例代码

当你使用一个 Hadoop 客户端的命令行界面（CLI）时，你可以将文件从本地复制到 HDFS 上，然后你可以通过下面的代码段查看前 10 行。

```
[hadoop@client-host ~]$ hadoop fs -ls /data
Found 4 items
drwxr-xr-x - hadoop supergroup 0 2012-07-12 08:55 /data/faa
-rw-r--r-- 1 hadoop supergroup 100 2012-08-02 13:29
/data/sample.txt
drwxr-xr-x - hadoop supergroup 0 2012-08-09 19:19 /data/wc
drwxr-xr-x - hadoop supergroup 0 2012-09-11 11:14 /data/weblogs

[hadoop@client-host ~]$ hadoop fs -ls /data/weblogs/

[hadoop@client-host ~]$ hadoop fs -mkdir /data/weblogs/in

[hadoop@client-host ~]$ hadoop fs -copyFromLocal
weblogs_Aug_2008.ORIG /data/weblogs/in

[hadoop@client-host ~]$ hadoop fs -ls /data/weblogs/in
Found 1 items
-rw-r--r-- 1 hadoop supergroup 9000 2012-09-11 11:15
/data/weblogs/in/weblogs_Aug_2008.ORIG

[hadoop@client-host ~]$ hadoop fs -cat
/data/weblogs/in/weblogs_Aug_2008.ORIG \
| head
10.254.0.51 - - [29/Aug/2008:12:29:13 -0700] "GGGG / HTTP/1.1"
200 1456
10.254.0.52 - - [29/Aug/2008:12:29:13 -0700] "GET / HTTP/1.1"
200 1456
```