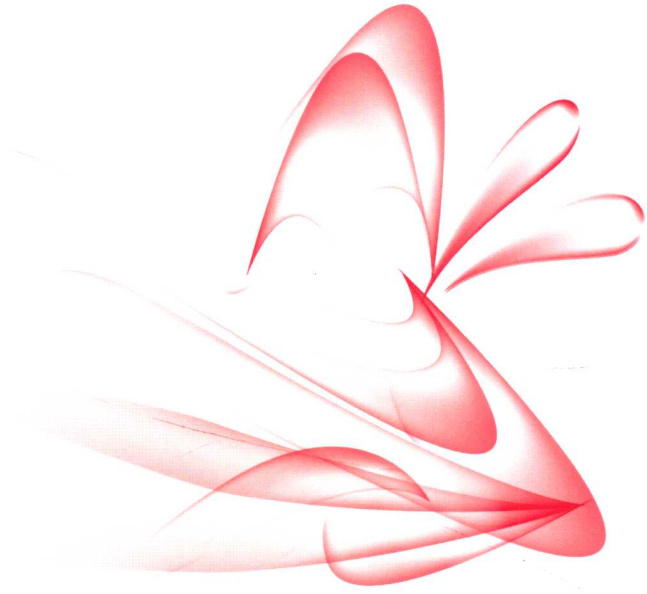
HZ BOOKS
华章科技

首部Hadoop YARN专著，资深Hadoop技术专家根据最新版本撰写，ChinaHadoop和51CTO等专业技术社区联袂推荐！

从应用角度系统讲解YARN的基本库和组件用法、应用程序设计方法、YARN上流行的各种计算框架，以及多个类YARN的开源资源管理系统。

从源代码角度深入分析YARN的设计理念与基本架构、各个组件的实现原理，以及各种计算框架的实现细节。

大数据
技术丛书



Hadoop Internals: in-depth study of YARN

Hadoop技术内幕

深入解析YARN架构设计与实现原理

董西成◎著

机械工业出版社
China Machine Press

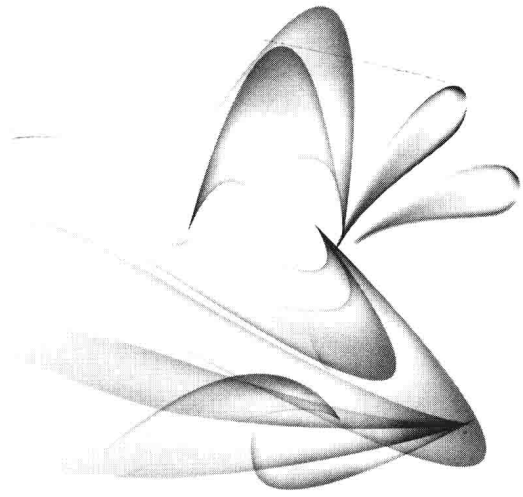
大数据技术丛书

Hadoop Internals: in-depth study of YARN

Hadoop技术内幕

深入解析YARN架构设计与实现原理

董西成◎著



机械工业出版社
China Machine Press

图书在版编目（CIP）数据

Hadoop 技术内幕：深入解析 YARN 架构设计与实现原理 / 董西成著 . —北京：机械工业出版社，2013.12（2014.1 重印）
（大数据技术丛书）

ISBN 978-7-111-44534-0

Ⅰ . H… Ⅱ . 董… Ⅲ . 数据处理软件 Ⅳ . TP274

中国版本图书馆 CIP 数据核字（2013）第 252913 号

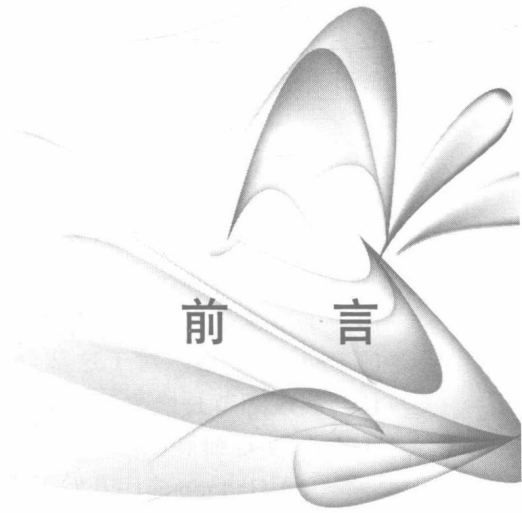
本书是“Hadoop 技术内幕”系列的第 3 本书，前面两本分别对 Common、HDFS 和 MapReduce 进行了深入分析和讲解，赢得了极好的口碑，Hadoop 领域几乎人手一册，本书则对 YARN 展开了深入的探讨，是首部关于 YARN 的专著。仍然由资深 Hadoop 技术专家董西成执笔，根据最新的 Hadoop 2.0 版本撰写，权威社区 ChinaHadoop 鼎力推荐。

本书从应用角度系统讲解了 YARN 的基本库和组件用法、应用程序设计方法、YARN 上流行的各种计算框架（MapReduce、Tez、Storm、Spark），以及多个类 YARN 的开源资源管理系统（Corona 和 Mesos）；从源代码角度深入分析 YARN 的设计理念与基本架构、各个组件的实现原理，以及各种计算框架的实现细节。

全书共四部分 13 章：第一部分（第 1~2 章）主要介绍了如何获取、阅读和调试 Hadoop 的源代码，以及 YARN 的设计思想、基本架构和工作流程；第二部分（第 3~7 章）结合源代码详细剖析和讲解了 YARN 的第三方开源库、底层通信库、服务库、事件库的基本使用和实现细节，详细讲解了 YARN 的应用程序设计方法，深入讲解和分析了 ResourceManager、资源调度器、NodeManager 等组件的实现细节；第三篇（第 8~10 章）则对离线计算框架 MapReduce、DAG 计算框架 Tez、实时计算框架 Storm 和内存计算框架 Spark 进行了详细的讲解；第四部分（第 11~13 章）首先对 Facebook Corona 和 Apache Mesos 进行了深入讲解，然后对 YARN 的发展趋势进行了展望。附录部分收录了 YARN 安装指南、YARN 配置参数以及 Hadoop Shell 命令等非常有用的资料。

机械工业出版社（北京市西城区百万庄大街 22 号　　邮政编码　100037）
责任编辑：孙海亮　　罗词亮
藁城市京瑞印刷有限公司印刷
2014 年 1 月第 1 版第 2 次印刷
186mm × 240 mm • 24.75 印张
标准书号：ISBN 978-7-111-44534-0
定　　价：69.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88378991　88361066　　投稿热线：（010）88379604
购书热线：（010）68326294　88379649　68995259　　读者信箱：hzjsj@hzbook.com



前　言

为什么要写这本书

在互联网巨头的带动下，开源软件 Hadoop 的应用变得越来越广泛，目前互联网、金融、银行、零售等行业均在使用或者尝试使用 Hadoop。IDC 对未来几年中国的预测中就专门提到了大数据，其认为未来几年，会有越来越多的企业级用户试水大数据平台和应用，而这之中，Hadoop 将成为最耀眼的“明星”。

尽管 Hadoop 整个生态系统是开源的，但由于它包含的软件种类过多，且版本升级过快，大部分公司，尤其是一些中小型公司，难以在有限的时间内快速掌握 Hadoop 蕴含的价值。此外，Hadoop 自身版本的多样化也给很多研发人员带来了很大的学习负担，尽管当前市面上已有很多参考书籍，但遗憾的是，能够深入剖析 Hadoop 内部实现细节的书籍少之又少，而本书则尝试弥补这一缺憾。本书是笔者继《Hadoop 技术内幕：深入解析 MapReduce 架构设计与实现原理》之后的又一本剖析 Hadoop 内幕的书籍。

本书介绍的 YARN（Yet Another Resource Negotiator）系统是 Hadoop 2.0 新增加的一个子项目（与 Common、MapReduce 和 HDFS 三个分支并列），它的引入使得分布式计算系统进入平台化时代，即各种计算框架可以运行在一个集群中，由资源管理系统进行统一管理和调度，它们共享整个集群中的资源进而提高资源利用率。

本书以 Hadoop 2.0 为基础，从基本概念、程序设计和内部实现等方面深入剖析了 Hadoop YARN。本书重点分析了 YARN 的核心实现以及运行在 YARN 上的计算框

架，其中，核心实现包括基础库、编程接口、ResourceManager 实现、资源调度器实现、NodeManager 实现等，而计算框架则包括离线计算框架 MapReduce、DAG 计算框架 Tez、实时计算框架 Storm 和内存计算框架 Spark 等。书中不仅详细介绍了 YARN 各个组件和计算框架的内部实现原理，而且结合源代码进行了深入剖析，使读者可以快速、全面地学习 Hadoop YARN 设计原理和实现细节。

读者对象

（1）Hadoop 二次开发人员

由于在扩展性、容错性和稳定性等方面的诸多优点，Hadoop 已被越来越多公司采用，而为了减少开发成本，大部分公司在 Hadoop 基础上进行二次开发，以打造属于公司内部的 Hadoop 平台。对于这部分 Hadoop 二次开发人员，深入而又全面地了解 Hadoop 的设计与实现细节是修改 Hadoop 内核的前提，而本书可帮助这部分读者快速而又全面地了解 Hadoop 实现细节。

（2）Hadoop 应用开发人员

如果要利用 Hadoop 进行高级应用开发，仅掌握 Hadoop 基本使用方法是远远不够的，必须对 Hadoop 框架的设计原理、架构和运作机制有一定的了解。对这部分读者而言，本书将带领他们全面了解 Hadoop 的设计和实现，加深对 Hadoop 框架的理解，提高开发水平，从而编写出更加高效的应用程序。

（3）Hadoop 运维工程师

对于一名合格的 Hadoop 运维工程师而言，适当地了解 Hadoop 框架的设计原理、架构和运作机制是十分有帮助的，这不仅可以更快地排除各种可能的 Hadoop 故障，也能够让 Hadoop 运维人员与研发人员进行更有效地沟通。通过阅读本书，Hadoop 运维人员可以了解到很多从其他书中无法获取的 Hadoop 实现细节。

（4）开源软件爱好者

Hadoop 是开源软件中的佼佼者，它在实现的过程中吸收了很多开源领域的优秀思想，同时也有很多值得学习的创新。尤为值得一提的是，本书分析 Hadoop 设计和实现的方式也许值得所有开源软件爱好者进行学习和借鉴。通过阅读本书，这部分读者不仅能领略到开源软件的优秀思想，还可以掌握分析开源软件源代码的方法和技巧，从而进一步提高使用开源软件的效率和质量。

如何阅读本书

本书分为四大部分（不包括附录）：

第一部分为基础篇（第 1 ~ 2 章），简单地介绍 Hadoop YARN 的环境搭建和基本设计架构，帮助读者了解一些基础背景知识。

第二部分为YARN核心设计篇（第3 ~ 7章），着重讲解YARN基本库、应用程序设计方法和运行时环境的实现，包括ResourceManager、NodeManager和资源调度器等关键组件的内部实现细节。

第三部分为计算框架篇（第8 ~ 10章），主要讲解当前比较流行的可运行在YARN上的计算框架，包括离线计算框架MapReduce、DAG计算框架Tez、实时计算框架Storm和内存计算框架Spark。

第四部分为高级篇（第11 ~ 13章），主要介绍了几个类似于Hadoop YARN的开源资源管理系统，包括Corona、Mesos等，并总结了资源管理系统的特点及发展趋势。

另外本书最后还添加了几个附录：附录A为YARN安装指南；附录B介绍了常见的YARN配置参数；附录C介绍了常用的Hadoop Shell命令；附录D为本书的所有参考资料，包括参考论文、Hadoop jira和网络资源等。

Hadoop YARN是Hadoop 2.0新引入的系统，对于大部分读者而言，该系统存在很多疑惑与未知之处，而本书正是尝试全方位剖析该系统。为了能够系统化地学习YARN，推荐读者从第1章的基础理论知识开始学习。

勘误和支持

由于笔者的水平有限，加之编写时间仓促，书中难免会出现一些错误或者不准确的地方，恳请读者批评指正。为此，笔者特意创建一个在线支持与应急方案的站点http://hadoop123.com。你可以将书中的错误发布在Bug勘误表页面中，同时如果你遇到任何问题，也可以访问Q&A页面，我将尽量在线上为读者提供最满意的解答。如果你有更多的宝贵意见，也欢迎发送邮件至邮箱dongxicheng@yahoo.com，期待能够得到你们的真挚反馈。

致谢

感谢我的导师廖华明副研究员，是她引我进入Hadoop世界。

感谢腾讯的蔡斌老师，正是由于他的推荐，才使得两本Hadoop书的出版成为可能。

感谢机械工业出版社华章公司的杨福川老师和孙海亮老师在这一年多的时间中始终支持我的写作，他们的鼓励和帮助使我顺利完成了本书。

感谢何鹏、姜冰、郑伟伟、战科宇、周礼、刘晏辰、王群等人给我提供的各种帮助。

最后感谢我的父母，感谢他们的养育之恩，感谢兄长的鼓励和支持，感谢他们时时刻刻给我信心和力量！感谢我的女朋友颛悦对我生活的细心照料与琐事上的宽容。

谨以此书献给我最亲爱的家人，以及众多热爱Hadoop的朋友们！

董西成

于北京

目　录

第三部分　计算框架篇

第四部分　高级篇

第一部分

准 备 篇

由于MRv1（MapReduce version 1）在扩展性、可靠性、资源利用率和多框架等方面存在明显不足，故Apache开始尝试对MapReduce进行升级改造，进而诞生了下一代MapReduce计算框架MRv2（MapReduce version 2）。由于MRv2将资源管理模块构建成了一个独立的通用系统YARN，这使得MRv2的核心从单一计算框架MapReduce转移为通用资源管理系统YARN。本书第一部分将介绍学习MRv2前的准备工作，并给出MRv2和YARN的基本概念和架构。