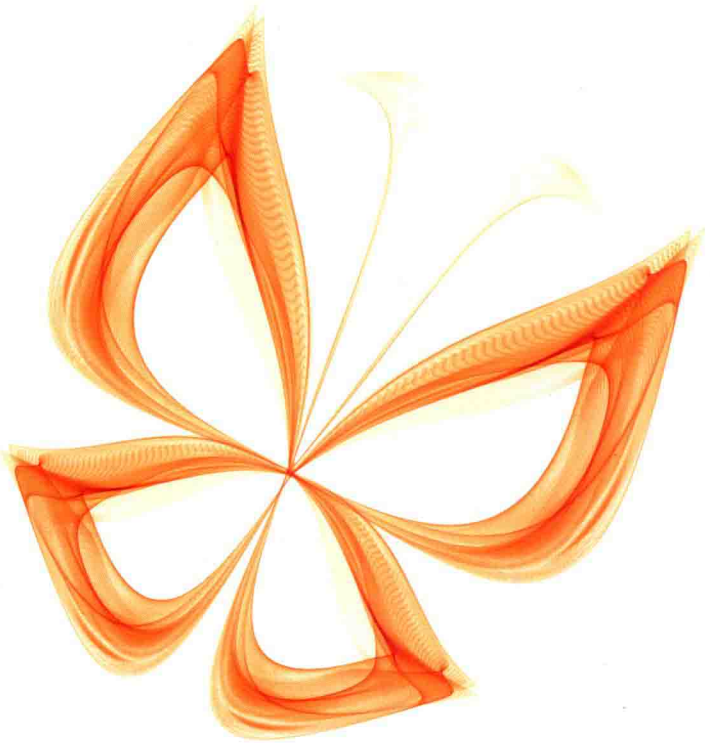
多位专家联袂推荐，360大数据专家撰写，基于Spark 2.1.0剖析架构与实现精髓

细化到方法级，提炼出多个流程图，立体呈现架构、环境、调度、存储、计算、部署、API七大核心设计

大数据技术丛书



The Art of Spark Kernel Design

Spark内核设计的艺术

架构设计与实现

耿嘉安◎著

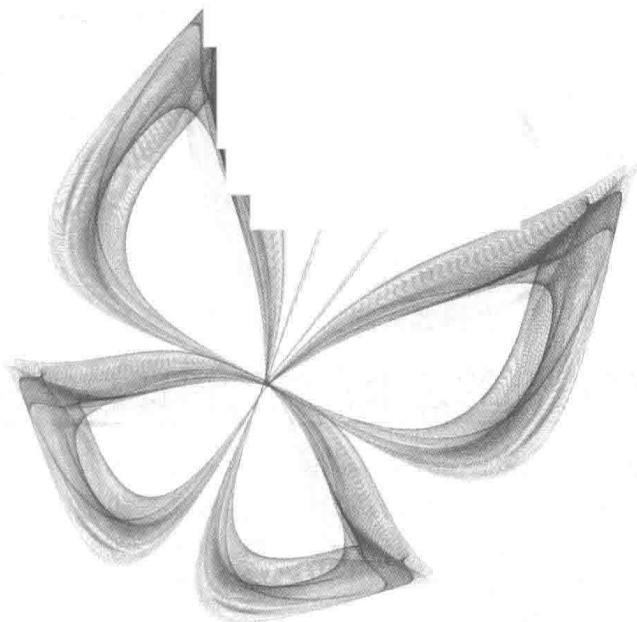
机械工业出版社
China Machine Press

大数据技术丛书

The Art of Spark Kernel Design

Spark内核设计的艺术
架构设计与实现

耿嘉安◎著



机械工业出版社
China Machine Press

图书在版编目（CIP）数据

Spark 内核设计的艺术：架构设计与实现 / 耿嘉安著 . —北京：机械工业出版社，2017.12
（大数据技术丛书）

ISBN 978-7-111-58439-1

I. S… II. 耿… III. 数据处理软件 IV. TP274

中国版本图书馆 CIP 数据核字（2017）第 278308 号

Spark 内核设计的艺术：架构设计与实现

出版发行：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：高婧雅　　责任校对：殷　虹
印　　刷：中国电影出版社印刷厂　　版　　次：2018 年 1 月第 1 版第 1 次印刷
开　　本：186mm × 240mm　1/16　　印　　张：44
书　　号：ISBN 978-7-111-58439-1　　定　　价：139.00 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）88379426　88361066　　投稿热线：（010）88379604
购书热线：（010）68326294　88379649　68995259　　读者信箱：hzit@hzbook.com

华章IT
HZBOOKS | Information Technology

Praise 本书赞誉

当年我在英国从事大数据工作，会经常去硅谷拜访大数据公司。其中最重要一个公司就是 Spark 创始人创建的 Databricks 了，最早一次是 2013 年 10 月，彼时 Databricks 刚起步，新办公室也尚在筹备。

4 年过去了，我们在大数据、流计算、图计算、分布式机器学习、深度学习等领域有了越来越多的高质量开源选择，但是 Spark 仍然是数据科学家们用得最多的工具之一，了解一点 Spark 底层技术的人都不得不对 Spark 的设计及其分布式计算的理论基础表示由衷敬佩。

本书对 Spark 内部高度抽象的数据结构 RDD、分布式 DAG 调度器 / 驱动器，以及高效的基于 Non-blocking IO 分布式计算框架 Akka/Netty 等内核设计进行了深度剖析，不可多得，是适合大型分布式计算架构师和资深开源贡献者阅读的参考书。

——蔡栋，万达网络科技集团总裁助理兼首席数据官、首席架构师

大数据技术生态其实是一个千姿百态的江湖。从学习技术的角度，最重要的是能将厚变薄，将纷繁复杂的信息进行归类和抽象。对应到大数据技术体系，虽然各种技术百花齐放，层出不穷，但大数据技术本质上无非解决 4 个核心问题：存储，计算，查询，挖掘。而 Spark 发展的短短几年，以迅雷不及掩耳之势推出 RDD、Spark Streaming、Spark SQL、GraphX、MLlib 等一系列模块，震撼了大数据圈。这本书结合了最新 Spark 2.x 版本，在设计思路和代码解析上做了很好的平衡，让开源代码爱好者，喜欢研究源码的同学汲取到一些阅读源码的方法。

——董飞，datatist 首席运营官、前 linkedin 资深工程师

初读本书有种似曾相识的感觉，Spark 还是那个 Spark，但是本书多了一些岁月的痕迹，在技术之上多了一些艺术，也更加注重读者的口味。大数据的书很多，能够写出艺术味道的不多，本书应该可以让你在大数据漫漫征途之中对价值多了一重思考，也可以让你在大数据之巅的惊天骇浪中多了一座灯塔。

——于俊，科大讯飞大数据专家

制度信息化，信息工具化，Spark 为大数据产业落地提供有力的技术支撑工具！它以内存计算为核心，以其通用、快速和完整的数据工具形成了一个强有竞争力的数据生态圈，成为大数据技术解决方案非常优秀的一个部分，越来越多企业应用部署 Spark。本书为那些想要成为一名合格的 Spark 工程师，或者致力于成为大数据行业的技术管理人才提供了很好的学习途径。相信读者只要掌握一门 Spark 技术，就能在大数据的海洋中遨游。感谢笔者为大数据产业做出的贡献！

——张涵诚，中关村大数据交易产业联盟副秘书长

本书对 Spark 原理的讲解与剖析都极具学习意义，作者细致分析了 Spark 源码的每一个关键细节，对初级用户及中高级用户都有指导意义。

——王欢，上海添锡信息技术有限公司技术总监

Preface 前 言

为什么写这本书

给本书写前言时，让我想起了两年前给《深入理解 Spark：核心思想与源码分析》[⊖]一书写前言的经历。我不禁想起崔护的《题都城南庄》这首诗，诗的内容是：

去年今日此门中，人面桃花相映红。

人面不知何处去，桃花依旧笑春风。

从核心思想和架构来看，Spark 依然是那个 Spark，但是我已经找了一个新的“东家”。我的年龄不知不觉中又长了两岁，Spark 也在大数据领域从“新贵”变成了“老人”。Spark 的版本从 0.x.x 到 2.x.x 基本上也是用了两年时间。

自从《深入理解 Spark：核心思想与源码分析》一书出版后，引起了一些市场反响，更难得的是得到了很多读者的反馈。一些热心的读者通过微信或者邮件向我指出了书中内容的很多不足之处，包括错别字、错误的描述、代码分析有点像流水账、提纲挈领的内容偏少、代码版本过低等。一些错误在修订的版本中得到了解决，有些修正的内容则通过单独写博客来补充。在与读者的沟通过程中，也纠正了我对一些问题的理解偏差。再次深深地感谢广大读者的支持与帮助！

一些读者对《深入理解 Spark：核心思想与源码分析》一书的内容非常肯定，希望能够出第 2 版，高婧雅编辑也一再“怂恿”我，但是我一直没有写第 2 版的打算。我当时希望有人能够以更好的方式写一本介绍和分析 Spark 2.0 版本的源码分析书籍，因为我感觉之前的写作方式的确不是很好。在我心中一直有个矛盾：如果源码太少，源码分析的书籍将退化成单纯讲原理的书籍，对于想深入理解 Spark 实现的读者来说这是不够的；如果源码太多，又让人有堆砌代码或者“混”篇幅的感觉。很多源码分析的书只是简单说说接口或者方法的功能，让人始终有种“雾里看花”的感觉。所以我一直很期待能有更好的方式来写作源码分析类的书。

在一年多的等待中，我始终没有发现类似书籍的出现，于是我打算再做一次尝试。这

⊖ 该书由机械工业出版社出版，书号 978-7-111-52234-8。

次摈弃了《深入理解 Spark：核心思想与源码分析》一书中按照代码执行流程分析的方式，改为先从整体上介绍一个系统，然后逐个分析每个组件的功能，最后将这些组件之间的关系用流程图的方式串联起来。本书的写作方式依然犯有代码过多的“毛病”，但我还是期待本书能带来一些新的气象。

本书的主要特色

- 按照源码分析的习惯设计，从脚本分析到初始化，再到核心内容。整个过程遵循由浅入深的基本思路。
- 每一章先对本章的内容有个总体介绍，然后深入分析各个组件的实现原理，最后将各个组件之间的关系通过执行流程来展现。
- 本书尽可能地用图来展示原理，以加速读者对内容的掌握。
- 本书讲解的很多实现及原理都值得借鉴，可以帮助读者提升架构设计、程序设计等方面的能力。
- 本书尽可能保留较多的源码，以便于初学者能够在脱离办公环境的地方（如地铁、公交等），也能轻松阅读。

读者对象

源码阅读是一项苦差事，人力和时间成本都很高，尤其对于刚刚接触 Spark 的人来说更是如此。本书尽可能保留源码，使得分析过程不至于产生跳跃感，目的是降低大多数人的学习门槛。如果你是从事 IT 工作 1 ～ 3 年的新人或者希望开始学习 Spark 的核心知识，本书非常适合你。如果你已经对 Spark 有所了解或者已经使用它，还想进一步提高自己，那么本书更适合你。如果你是一个开发新手，对 Java、Linux 等基础知识还不是很了解的话，本书可能不太适合你。如果你已经对 Spark 有深入的研究，本书也许可以作为你的参考资料。

总体来说，本书适合以下人群：

- 已经了解过 Spark，但还想深入理解 Spark 实现原理的人；
- 大数据技术爱好者；
- 对性能优化和部署方案感兴趣的运维工程师与架构师；
- 开源代码爱好者，喜欢研究源码的同学可以通过本书学到一些阅读源码的方式、方法。

本书不会教你如何开发 Spark 应用程序，而只拿 word count 的经典例子做演示。本书会简单介绍 Hadoop MapReduce、Hadoop YARN、Mesos、Alluxio（Tachyon）、ZooKeeper、HDFS、Akka、Jetty、Netty，但不会过多介绍这些框架的使用，因为市场上已经有丰富的书籍供读者挑选。本书也不会过多介绍 Scala、Java、Shell 的语法，读者可以在市场上选择适合自己的书籍阅读。本书将无比适合那些想要破解“潘多拉魔盒”的人！

如何阅读本书

本书一共有 10 章内容，主要包括以下部分。

准备部分（第 1～2 章）：简单介绍了 Spark 的环境搭建和基本原理，帮助读者了解一些背景知识。

基础部分（第 3～5 章）：介绍 Spark 的基础设施、SparkContext 的初始化、Spark 执行环境等内容。

核心部分（第 6～9 章）：这是 Spark 最为核心的部分，包括存储体系、调度系统、计算引擎、部署模式等。

API 部分（第 10 章）：这部分主要对 Spark 的新旧 API 进行对比，对新 API 进行介绍。

本书最后的附录中还包括一些内容：附录 A 介绍的是 Spark 中最常用的工具类 Utils；附录 B 是 Akka 的简介；附录 C 为 Jetty 的简介和工具类 JettyUtils 的介绍；附录 D 为 Metrics 库的简介和 Metrics 中部分 API 的介绍；附录 E 演示了 Hadoop 1.0 版本中的 word count 例子；附录 F 介绍了工具类 CommandUtils 的常用方法；附录 G 是关于 Netty 的简介和工具类 NettyUtils 的介绍；附录 H 是对 Spark 中的 RPC 工具类 RpcUtils 的介绍。

为了降低读者阅读理解 Spark 源码的门槛，本书尽可能保留源码实现。本书以 Spark 2.1.0 版本为主，有兴趣的读者也可按照本书的方式，阅读 Spark 的最新源码。

勘误

本书内容很多，限于笔者水平有限，书中内容难免有错误之处。如果你对本书有任何问题或者意见，都可以通过邮箱 beliefer@163.com 或者博客 http://blog.csdn.net/beliefer 联系我，给我提交你的建议或者想法，我将怀着一颗谦卑之心与大家共同进步。

致谢

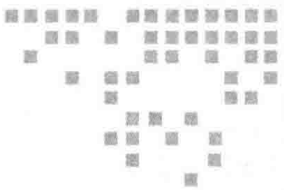
感谢我们生活在信息时代，让我们有机会接触互联网与大数据；感谢父母多年来在学习、工作及生活上的帮助与支持；感谢妻子在生活中的照顾和谦让。

感谢高婧雅编辑给予本书出版的大力支持与帮助。

感谢我在大数据路上的领路人——和仲；感谢热衷于技术的王欢对本书内容提出的宝贵建议；感谢对本书内容进行审阅的余尧尧和马晓波；感谢对本书内容有过帮助的读者朋友们。

耿嘉安

目　录 Contents



第1章 Chapter 1
环境准备

“凡事豫则立，不豫则废。言前定，则不跲；事前定，则不困。”

——《礼记·中庸》

本章导读

学习一个工具的最好途径，就是使用它。这就好比《极品飞车》玩得好的同学，未必真的会开车，要学习车的驾驶技能，就必须用手触摸方向盘，用脚感受刹车与油门的力道。在IT领域，在深入了解一个系统的原理和实现细节之前，应当先准备好它的运行环境或者源码阅读环境。如果能在实际环境下安装和运行Spark，显然能够提升读者对Spark的一些感受，对系统能有个大体的印象，有经验的工程师甚至能够猜出一些Spark在实现过程中采用的设计模式和编程模型。

当你通过一些途径知道了系统的原理之后，难道不会问问自己：这是怎么做到的？如果只是游走于系统使用、原理了解的层面，是永远不可能真正理解整个系统的。很多IDE本身带有调试的功能，每当你阅读源码，陷入重围时，调试能让我们更加理解运行期的系统。如果没有调试功能，不敢想象阅读源码的难度有多大。

本章旨在帮助读者搭建基本的运行环境、构建源码学习环境。主要包括以下内容。

- 在Linux中搭建基本的Spark执行环境。
- spark-shell的基本使用。
- spark-shell的源码剖析。
- 在Mac OS中搭建Spark源码阅读环境。

1.1 运行环境准备

考虑到大部分公司开发和生产环境都采用 Linux 操作系统，所以笔者选用了 64 位的 Linux。在正式安装 Spark 之前，先要找台好机器。为什么？因为笔者在安装、编译、调试的过程中发现 Spark 非常耗费内存，如果机器配置太低，恐怕会运行不起来。Spark 的开发语言是 Scala，而 Scala 需要运行在 JVM 之上，因而搭建 Spark 的运行环境应该包括 JDK 和 Scala。

1.1.1 安装 JDK

自 Spark 2.0.0 版本开始，Spark 已经放弃了对 Java 7 的支持，所以需要选择 Java 8。我们还需要使用命令 getconf LONG_BIT 查看 Linux 机器是 32 位还是 64 位，然后下载相应版本的 JDK 并安装。

下载地址：http://www.oracle.com/technetwork/java/javase/downloads/index.html。

配置环境：

```
cd ~
vim .bash_profile
```

添加如下配置：

```
export JAVA_HOME=/opt/java
export PATH=$PATH:$JAVA_HOME/bin
export CLASSPATH=.:$JAVA_HOME/lib/dt.jar:$JAVA_HOME/lib/tools.jar
```

输入以下命令使环境变量快速生效：

```
source .bash_profile
```

安装完毕后，使用 java –version 命令查看，确认安装正常，如图 1-1 所示。

```
EBJ1266:~ user$ java -version
java version "1.8.0_102"
Java(TM) SE Runtime Environment (build 1.8.0_102-b14)
Java HotSpot(TM) 64-Bit Server VM (build 25.102-b14, mixed mode)
```

图 1-1 查看 java 安装是否正常

1.1.2 安装 Scala

由于从 Spark 2.0.0 开始，Spark 默认使用 Scala 2.11 来编译、打包，不再是以前的 Scala 2.10，所以我们需要下载 Scala 2.11。

下载地址：http://www.scala-lang.org/download/。

选择 Scala 2.11 的版本进行下载，下载方法如下：

```
wget https://downloads.lightbend.com/scala/2.11.8/scala-2.11.8.tgz
```

移动到选好的安装目录，例如：