由Elasticsearch内核开发工程师编写，从源码和设计角度分析Elasticsearch的内部原理，为合理、高效地使用Elasticsearch提供理论指导，并为大规模应用和维护过程中的常见问题提供具体的优化措施和故障诊断方法

Broadview
www.broadview.com.cn

Elasticsearch
源码解析与优化实战

张超 / 著

中国工信出版集团
電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

Elasticsearch
源码解析与优化实战

张超 / 著

電子工業出版社
Publishing House of Electronics Industry
北京•BEIJING

内 容 简 介

本书介绍了 Elasticsearch 的系统原理，旨在帮助读者了解其内部原理、设计思想，以及在生产环境中如何正确地部署、优化系统。系统原理分两方面介绍，一方面详细介绍主要流程，例如启动流程、选主流程、恢复流程；另一方面介绍各重要模块的实现，以及模块之间的关系，例如 gateway 模块、allocation 模块等。本书的最后一部分介绍如何优化写入速度、搜索速度等大家关心的实际问题，并提供了一些诊断问题的方法和工具供读者参考。

本书适合对 Elasticsearch 进行改进的研发人员、平台运维人员，对分布式搜索感兴趣的朋友，以及在使用 Elasticsearch 过程中遇到问题的人们。

图书在版编目（CIP）数据

Elasticsearch 源码解析与优化实战 / 张超著. —北京：电子工业出版社，2018.11
ISBN 978-7-121-35216-4

Ⅰ. ①E… Ⅱ. ①张… Ⅲ. ①搜索引擎—程序设计 Ⅳ. ①TP391.3

中国版本图书馆 CIP 数据核字（2018）第 238923 号

责任编辑：陈晓猛
印　　刷：三河市良远印务有限公司
装　　订：三河市良远印务有限公司
出版发行：电子工业出版社
　　北京市海淀区万寿路 173 信箱　　邮编：100036
开　　本：787×980　1/16　　印张：22.5　　字数：432 千字
版　　次：2018 年 11 月第 1 版
印　　次：2018 年 11 月第 1 次印刷
定　　价：89.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：010-51260888-819，faq@phei.com.cn。

前言

我们可以在不关心原理的情况下使用 Elasticsearch（以下简称 ES），但要想用好 ES，就必须熟知其内部原理。

为什么要阅读代码？在传统软件行业，技术文档非常丰富。当开展一个项目时，从需求分析，到概要设计、详细设计，每个步骤都有相应的文档，从项目的整体架构、技术方案选型，到流程图、类图，细化到每个接口及参数。在这种情况下，想要搞清楚系统原理，并不需要阅读代码，文档上什么都有。但是互联网产品迭代快，技术文档不全，想要搞清楚原理，只能阅读代码，相当于从代码中逆向理解设计思想。

通过分析源码，我们可以有以下收获：

理解设计思想 当我们面临要解决的问题或实现的目标时，往往有多种方案可以选择。无论表面上看起来多么简单的架构，其背后都经过了深思熟虑。思考一下为什么使用现在的方案？有没有更好的解决方案？

探究内部机制的原理 某个技术点是怎么实现的？

搞明白执行流程 某个过程是什么样的，都做了什么？有几步？先做什么，后做什么？

熟悉代码结构 如果需要进行二次开发，则给出代码入口和调用关系，有时候找到某个逻辑的代码实现要花很多时间。

学以致用 借鉴其设计理念，掌握其解决问题的方式和方法，将来面对类似的问题时可以参考。

本书结构

本书由四部分组成，第一部分为基础知识和环境准备（第 1～2 章）；第二部分介绍 ES 的主要流程（第 3～10 章），包括集群启动流程、节点启动/关闭流程、选主流程、读写流程、搜索流程和索引恢复流程；第三部分主要介绍重要内部模块（第 11～17 章），包括 gateway 模块、allocation 模块、Snapshot 模块、Cluster 模块、Transport 模块和 ThreadPool 模块等；第四部分介

绍优化和诊断方法（第 18～22 章），包括写入速度优化、搜索速度优化、磁盘使用量优化，以及在生产环境中的实际应用建议，第 22 章介绍常用的问题诊断方法，排查集群遇到的问题。

术语约定

ES 中有一些特有的概念，这些概念对应的中文翻译约定如下：

- 分片（shard）；
- 主分片（primary shard），简称 P；
- 分片副本（特指数据的一个分片，无论主分片，还是副分片）；
- 副分片（replica shard），简称 R；
- 分片分配（shard allocation）；
- 集群状态（cluster state）；
- 分配决策（allocation decision）；
- 分配感知（allocation awareness）；
- 分配标识（allocation IDs）；
- 追踪（tracking）；
- 事务日志（translog）；
- 同步集合（in-sync set）。

行文约定

虽然本书是一本源码分析类图书，但原则上尽量少贴代码，引用的代码只是为了说明原理，因此所引用的代码并不保证和源码完全一致，对非核心逻辑有所删减，同时在代码块中，函数参数可能被省略，省略的函数参数用“…”表示，如：

```
executeBulk(...);
```

在引用代码中的某个方法时，使用#号分隔类名与方法名：

```
类名#方法名
```

一个索引由许多分片组成。我们用如下方式表示索引 website 的第 0 个分片：

```
website[0]
```

联系

读者有任何意见和建议都可以联系作者，邮箱：elasticsearchbook@163.com。

本书配套网站：www.elasticsearchbook.cn。

致谢

感谢李欣杰和郭东东，他们带我走进搜索领域；感谢韩洪伟，他让我学到了很多搜索系统的知识。欣杰和老韩都是资深的搜索架构师，能够和优秀的团队共事是我的荣幸。感谢 ES 团队的同事段军义，我们互相学习，一起解决了很多麻烦的问题。感谢出版社的策划编辑陈晓猛先生，他为本书的写作提供了很多建设性意见，并且耐心地编校了本书，让本书得以顺利出版。

感谢我的妻子和三岁的女儿，我爱你们！

张超

读者服务

轻松注册成为博文视点社区用户（www.broadview.com.cn），扫码直达本书页面。

- **提交勘误**：您对书中内容的修改意见可在 提交勘误 处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **交流互动**：在页面下方 读者评论 处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：http://www.broadview.com.cn/35216

目录