

# 《分布式实时计算框架原理及实践案例》

## 图书基本信息

书名：《分布式实时计算框架原理及实践案例》

13位ISBN编号：9787121296209

出版时间：2016-9-1

作者：王成光

页数：292

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：[www.tushu111.com](http://www.tushu111.com)

# 《分布式实时计算框架原理及实践案馈

## 内容概要

“授人以鱼不如授人以渔”，《分布式实时计算框架原理及实践案例》是作者以如此初心写成的，主要参考当前主流分布式实时计算框架Storm的任务分发和Spark Streaming的Mini-Batch设计思想，以及底层实现技术，开源了作者自研的轻量级分布式实时计算框架——Light\_drtc，并且重点介绍设计思想和相关实现技术（Kafka/RabbitMQ、Redis/SSDB、GuavaCache、MongoDB、HBase、ES / SolrCloud、Thrift、Avro、Jetty），后从工程角度向大家介绍完整的个性化推荐系统，并实例介绍Light\_drtc在用户画像实时更新的应用。《分布式实时计算框架原理及实践案例》描述浅显易懂，希望读者理解分布式实时计算的实现原理，并快速上手解决实际问题。

# 《分布式实时计算框架原理及实践案馈

## 作者简介

王成光，目前任职一点资讯研发架构师，硕士毕业8年，曾先后任职窝窝团、优购、搜狐、网易等架构师、技术专家职位，专注于搜索、推荐、数据挖掘领域研发工作，涉足技术范围：

搜索：ES / SolrCloud

分布式计算：Hadoop、Storm和Spark

MQ：Kafka、RabbitMQ、ActiveMQ、ZeroMQ

NoSQL：Redis/SSDB、Mongo3.0、HBase1.0、AeroSpike

SOA微服务：RPC和Web Service

## 书籍目录

第1章 分布式实时计算框架介绍	1
1.1 分布式计算Hadoop	1
1.2 分布式实时计算	3
1.2.1 Spark Streaming	3
1.2.2 Storm	6
1.2.3 其他框架	8
1.3 为什么自研	8
1.4 总结	10
第2章 light_drtc 简介及使用说明	11
2.1 light_drtc 框架简介	11
2.2 light_drtc 代码结构	12
2.3 light_drtc 重要配置项	14
2.4 light_drtc 和Storm 比较	15
2.5 light_drtc 使用说明	16
2.5.1 ACN ( AN 和CN 整合 ) 作为独立服务	16
2.5.2 CN、AN 作为独立服务	20
2.5.3 任务计算JN	23
2.6 总结	26
第3章 light_drtc 核心技术实现	27
3.1 light_drtc 技术架构	27
3.2 light_drtc 计算框架设计思想	30
3.2.1 CN 设计思想	30
3.2.2 AN 多主模式设计思想	31
3.2.3 JN 设计思想	34
3.3 light_drtc 核心技术的实现	36
3.3.1 实时收集数据CN	36
3.3.2 任务协调管理AN	40
3.3.3 任务计算JN	49
3.4 总结	50
第4章 消息队列MQ	51
4.1 消息队列使用场景	51
4.2 消息队列原理	53
4.2.1 MQ 使用流程	53
4.2.2 MQ 基本概念	54
4.2.3 MQ 通信模式	55
4.2.4 目前知名MQ 比较	56
4.3 MQ 消费状态监控	61
4.3.1 KafkaOffsetMonitor 介绍	62
4.3.2 KafkaOffsetMonitor 部署	62
4.4 RabbitMQ 和Kafka 的基本使用	64
4.4.1 RabbitMQ 读写实例	64
4.4.2 Kafka 读写实例	68
4.5 总结	71
第5章 内存数据库Redis3.0 及SSDB	72
5.1 Redis 相关介绍	72
5.1.1 Redis3.0 集群架构	73
5.1.2 Redis3.0 集群选举与容错	74

5.1.3 SSDB 简介.....75	
5.2 Redis3.0 集群搭建..76	
5.2.1 集群所依赖的Ruby 环境.....77	
5.2.2 Redis 集群创建77	
5.2.3 Redis 集群验证78	
5.2.4 SSDB 简单部署.....79	
5.3 Redis 管理及使用...81	
5.3.1 Redis 基本使用81	
5.3.2 Redis 管理.....83	
5.4 Redis 客户端应用...86	
5.4.1 Redis3.0 客户端.....86	
5.4.2 SSDB 客户端...89	
5.5 本地缓存Guava Cache.....90	
5.5.1 认识Guava Cache .....90	
5.5.2 Guava Cache 使用.....91	
5.5.3 Java 客户端使用.....94	
5.6 总结.....97	
第6章 NoSQL : MongoDB3.0 和HBase1.0 .....98	
6.1 MongoDB3.0 和HBase1.0 新特性.....99	
6.1.1 MongoDB3.0 新特性..99	
6.1.2 HBase1.0 新特性.....102	
6.1.3 MongoDB 和HBase 比较....104	
6.2 MongoDB3.0 集群和索引.....105	
6.2.1 MongoDB3.0 集群....105	
6.2.2 Mongo 索引介绍.....107	
6.3 HBase 底层实现介绍.....108	
6.3.1 HBase 相关Hadoop 体系....108	
6.3.2 HBase 系统架构..... 110	
6.4 Mongo 和HBase 客户端使用. 113	
6.4.1 Mongo 客户端 113	
6.4.2 HBase 客户端. 119	
6.5 总结.....124	
第7章 全文检索 : ElasticSearch2.x..125	
7.1 认识ElasticSearch 和Solr.....125	
7.1.1 ElasticSearch 和Solr 基本介绍.....125	
7.1.2 ES 基本概念...127	
7.1.3 ES 和SolrCloud 集群结构...129	
7.1.4 ES 使用案例...130	
7.2 ES 和Solr 比较分析.....131	
7.2.1 ES 和Solr 发展比较.131	
7.2.2 ES 和Solr 综合比较.132	
7.3 ES 集群介绍135	
7.3.1 插件安装.....135	
7.3.2 中文分词安装136	
7.3.3 ES2.X 集群节点类型138	
7.3.4 ES 配置事项...142	
7.4 ES 客户端使用.....144	
7.4.1 ES 客户端连接.....145	
7.4.2 ES 基本操作...146	

7.4.3 ES 高级使用...	150
7.5 ES 在自研框架中的作用.....	154
7.6 总结.....	155
第8章 微服务架构通信——RPC 和Web Service .....	156
8.1 微服务架构由来.....	156
8.1.1 微服务与SOA 比较..	157
8.1.2 微服务架构的优缺点	159
8.1.3 微服务雪崩效应的防范.....	161
8.2 RPC 介绍及实践...	163
8.2.1 Thrift/Nifty 介绍.....	163
8.2.2 Avro 介绍.....	168
8.2.3 Dubbo/Dubbox 介绍..	180
8.2.4 GRPC/ProtoBuf 介绍	185
8.2.5 ZeroC ICE.....	191
8.3 Web Service 介绍及实践.....	199
8.3.1 SOAP 和Rest .	200
8.3.2 JWS ( JDK 自身实现Web Service ) .....	202
8.3.3 Jetty : 嵌入式Servlet 容器..	204
8.3.4 基于Spring MVC.....	206
8.3.5 其他Web Service 框架.....	211
8.4 总结.....	212
第9章 综合实例：新闻推荐中的用户画像近实时更新	213
9.1 个性化推荐系统组成.....	213
9.1.1 用户行为收集	214
9.1.2 行为日志解析	216
9.1.3 常用推荐算法	221
9.1.4 用户画像数据仓库...	245
9.1.5 元数据索引库	247
9.1.6 用户推荐服务	248
9.2 新闻推荐中用户画像近实时更新设计.....	248
9.2.1 新闻推荐中用户画像构成...	250
9.2.2 新闻推荐中用户画像标签数据字典.....	251
9.2.3 新闻推荐用户画像实时更新流程..	257
9.3 新闻推荐用户画像近实时更新技术实现..	260
9.3.1 Storm 接入Kafka 实时计算实例...	260
9.3.2 Spark Streaming 接入Kafka 实时计算实例.....	265
9.3.3 Light_drtc 接入Kafka.....	270
9.3.4 用户画像实时更新核心实现.....	270
9.4 总结.....	280

# 《分布式实时计算框架原理及实践案馈

## 精彩短评

- 1、垃圾书，千万别买！！！！贴代码，抄简介，还不如看博客和各大技术官网！！！！骗子书！！太坑！他妹的！！幸好是特价买的，太恶心了！！！！
- 2、本书作者向读者朋友介绍其所独立研发的轻量级分布式实时计算框架light\_drtc的设计思想和核心实现原理，并向读者介绍一个完整的推荐系统所需要相关技术：消息MQ、缓存DB、NoSQL、全文检索以及常用各种微服务技术实现框架，对比分析，每个组件都有相关实践案例，让读者更加清晰如何在实际中进行技术选型，最后向读者介绍一个完整的个性化推荐系统组成、相关算法和用户画像实时更新详细设计和核心代码实现。本书偏工程化，强调实战经验。
- 3、这书。。。。
- 4、骗子写的骗钱的书，目录看上去丰富，每一章都泛泛而谈，还不如网上的博客

# 《分布式实时计算框架原理及实践案馈

## 版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:[www.tushu111.com](http://www.tushu111.com)