



[大数据技术丛书_下载链接1](#)

著者:吉奥兹 (P.Taylor Goetz)

出版者:机械工业出版社

出版时间:2015-1-1

装帧:平装

isbn:9787111484387

《Storm分布式实时计算模式》由Apache Storm项目核心贡献者吉奥兹、奥尼尔亲笔撰写，融合了作者丰富的Storm实战经验，通过大量

示例，全面而系统地讲解使用Storm进行分布式实时计算的核心概念及应用，并针对不同的应用场景，给出多种基于Storm的设计模式，为读者快速掌握Storms分布式实时计算提供系统实践指南。

《Storm分布式实时计算模式》分为10章：第1章介绍使用storm建立一个分布式流式计算应用所涉及的核心概念，包括storm的数据结构、开发环境的搭建，以及Storm程序的开发和调试技术等；第2章详细讲解storm集群环境的安装和搭建，以及如何将topology部署到分布式环境中；第3章通过传感器数据实例详细介绍Trident topology；第4章讲解如何使用Storm和Trident进行实时趋势分析；第5章介绍如何使用Storm进行图形分析，将数据持久化存储在图形数据库中，通过查询数据来发现其中潜在的联系；第6章讲解如何在Storm上使用递归实现一个典型的人工智能算法；第7章演示集成Storm和非事务型系统的复杂性，通过集成Storm和开源探索性分析架构Druid实现一个可配置的实时系统来分析金融事件。

第8章探讨Lambda体系结构的实现方法，讲解如何将批处理机制和实时处理引擎结合起来构建一个可纠错的分析系统；第9章讲解如何将Pig脚本转化为topology，并且使用Storm-YARN部署topology，从而将批处理系统转化为实时系统；第10章介绍如何在云服务提供商提供的主机环境下部署和运行 Storm。

作者介绍:

目录: 前言

作者简介

第1章 分布式单词计数

1.1 Storm topology的组成部分——stream、spout和bolt

1.1.1 Storm

1.1.2 spout

1.1.3 bolt

1.2 单词计数topology的数据流

1.2.1 语句生成bolt

1.2.2 语句分割bolt

1.2.3 单词计割bolt

1.2.4 上报bolt

1.3 实现单词计数top

1.3.1 配置开发环境

1.3.2 实现Sentence

1.3.3 实现语句分割bolt

1.3.4 实现单词计割bolt

1.3.5 实现上报bolt

1.3.6 实现单词计数topo

1.4 Storm的并发机制

1.4.1 WordCountTopology的并发机制

1.4.2 给topology增加woker

1.4.3 配置executor和task

1.5 理解数据流分组
1.6 有保障机制的数据处理
1.6.1 spout的可靠性
1.6.2 bolt的可靠性
1.6.3 可靠的单词计数
总结
第2章 配置Storm集群
2.1 Storm集群的框架
2.1.1 理解nimbus守护进程
2.1.2 supervisor守护进程的工作方式
2.1.3 Apache ZooKeeper简介
.....
第3章 Trident和传感器数据
第4章 实时趋势分析
第5章 实时图形分析
第6章 人工智能
第7章 整合Druid进行金融分析
第8章 自然语言处理
第9章 在Hadoop上部署Storm进行广告分析
第10章 云环境下的S
• • • • • ([收起](#))

[大数据技术丛书 下载链接1](#)

标签

Storm

大数据

实时计算

计算机

数据平台

软件开发

编程

江南白衣推荐

评论

只看了前面几章，算作了解storm的工作原理。后面的具体几个应用实例都没有看，因为了解原理之后这些相当于应用吧，打算进一步看看相关的源码了。

看了1, 2, 3章，还是看看文档，写点demo，然后实际项目中精进。
Storm学习系列，持续更新。<https://github.com/vonzhou/storm-examples>

看了第一章，了解Storm如何消费数据和怎么进行计算的之后就可以跟着公司里前辈写的代码撸起来了把。

基本上前面三章就可以干活了，但是后面几章虽然例子一般，但是也有有价值的东西。

这本书看起来有些鸡肋的感觉，前面70页Storm本身的讲解还算清楚。本是冲着书目中后面丰富的例子去的，但实际看完，并未觉得有多少有价值的内容留下，不尽如人意。说到底，Storm本身也是不够吸引人，将来应该很有更棒的流数据处理软件。

可能是我对分布式实时流计算接触的不多罢，读起来觉得挺好，清晰干净，书不厚，作为入门还是可以的

storm解决案例，适合初学者，比较浅，不够深入 想知道storm Internal的，还是省点钱吧

前两章例子适合入门，后边章节适合开拓思路。

相比国内的几本Storm书，算写的不错的了，每个例子代码都非常完整，可以运行

[大数据技术丛书 下载链接1](#)

书评

[大数据技术丛书 下载链接1](#)