

Alibaba Group | 技术丛书
阿里巴巴集团

Broadview®
www.broadview.com.cn

实时

大数据丛书 | 十二五国家重点图书出版规划项目

海量

Storm 实战

构建大数据实时计算

高效

|| 阿里巴巴集团数据平台事业部商家数据业务部 编著 ||

可靠

storm

电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

 **Alibaba Group** | 技术丛书
阿里巴巴集团

大数据丛书 | 十二五国家重点图书出版规划项目

Storm 实战

构建大数据实时计算

|| 阿里巴巴集团数据平台事业部商家数据业务部 编著 ||

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书是一本系统并且具有实践指导意义的Storm工具书和参考书，对Storm整个技术体系进行了全面的讲解，不仅包括对基本概念、特性的介绍，也涵盖了一些原理说明。本书的实战性很强，各章节都提供了一些小案例，同时对于本地，以及集群环境的部署有详细介绍，易于理解，操作性强。

全书一共10章：第1章全面介绍了Storm的特性、能解决什么问题，以及和其他流计算系统的对比；第2章通过实际运行一个简单的例子，以及介绍本地环境和集群环境的搭建，让读者对Storm有了直观的认识；第3章深入讲解了Storm的基本概念，同时实现一个Topology运行；第4章和第5章阐述了Storm的并发度、可靠处理的特性；第6章~第8章详细而系统地讲解了几个高级特性：事务、DRPC和Trident；第9章以实例的方式讲解了Storm在实际业务场景中的应用；第10章总结了几个在大数据场景应用过程中遇到的经典问题，以及详细的排查过程。

本书既适合没有Storm基础的初学者系统地学习，又适合有一定Storm基础但是缺乏实践经验的读者实践和参考。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目(CIP)数据

Storm 实战：构建大数据实时计算 / 阿里巴巴集团数据平台事业部商家数据业务部编著. —北京：电子工业出版社，2014.8

(大数据丛书. 阿里巴巴集团技术丛书)

ISBN 978-7-121-22649-6

I. ①S… II. ①阿… III. ①数据处理 IV. ①TP274

中国版本图书馆 CIP 数据核字(2014)第 050578 号

策划编辑：刘 皎

责任编辑：徐津平

印 刷：三河市鑫金马印装有限公司

装 订：三河市鑫金马印装有限公司

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：900×1280 1/32 印张：5.75 字数：91千字

版 次：2014年8月第1版

印 次：2014年8月第1次印刷

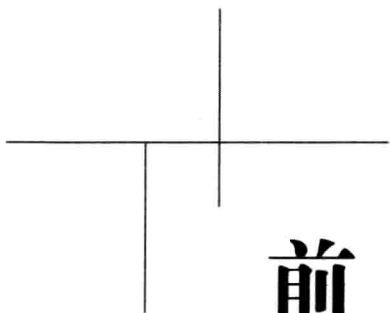
印 数：4000册 定价：59.00元



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。

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

服务热线：(010) 88258888。



前言

实时流计算

互联网从诞生的一刻起，对世界的最大改变就是让信息能够实时交互，从而大大提高各个环节的效率。正因为大家对信息实时响应、实时交互的需求，软件行业除了个人操作系统之外，数据库（更精确地说是关系型数据库）应该是软件行业发展最快、收益最为丰厚的产品了。记得 20 世纪 90 年代，很多银行别说实时转账，连实时查询都做不到，但是数据库和高速网络改变了这个情况。

互联网的进一步发展，从 Portal 信息浏览型到 Search 信息搜索型再到 SNS 关系交互传递型，以及电子商务、互联网旅游生活产品等，将人们生活中的流通环节在线化。对效率的要求让大家对实时性的要求进一步提升，而信息的交互和沟通正在从点对点向信息链甚至信息网的方向发展，这样必然带来数据在各个维度的交叉关联，数据爆炸已不可避免。因此流式处理加 NoSQL 产品应运而生，分别解决实时处理框架和数据大规模存储计算的问题。

早在 2000 年初，诸如 UC 伯克利、斯坦福等大学就开始了对流式数据处理的研究，但是由于更多地关注于金融行业的业务场景或者互联网流量监控的业务场景，以及当时互联网数据场景的限制，造成了研究多是基于对传统数据库处理的流式化，对流式框架本身的研究偏少。目前这样的研究逐渐没有了声音，工业界将更多的精力转向了实时数据库。

2010 年 Yahoo! 对 S4 的开源，2011 年 Twitter 对 Storm 的开源，改变了这个现状。以前互联网的开发人员在做一个实时应用的时候，除了要关注应用逻辑计算处理本身，还要

为数据的实时流转、交互、分布大伤脑筋。现在的情况却大为不同，以 **Storm** 为例，开发人员可以快速搭建一套健壮、易用的实时流处理框架，配合 **SQL** 产品或者 **NoSQL** 产品或者 **MapReduce** 计算平台，就可以以低成本做出很多以前很难想象的实时产品，比如量子恒道品牌旗下的多个产品就是构建在 **Storm** 实时流处理平台上的。

本书是一本对 **Storm** 的基础介绍手册，但是我们也希望它不仅仅是一本 **Storm** 的使用手册，我们会在其中加入更多在实际数据生产过程中的经验和应用架构，最终的目的是帮助所有愿意使用实时流处理框架的技术同仁，同时也默默地改变这个世界。

在本书即将出版之际，**Storm** 已经成功发布了 **0.9.0** 版本，追加了一些新的特性，如使用 **Netty** 作为新的消息传输层、提供日志查看 **UI** 等，同时修复了大量跟稳定性相关的 **BUG**。本次发布对茁壮成长的 **Storm** 来说是一次巨大的进步。新版本的 **Storm** 在系统结构及使用方式方面，并没有太多变化，本书可以帮助你快速掌握应用 **Storm** 的知识和技能。

本书由团队中多位同学先后参与合作完成，为体现阿里巴巴的文化，这里列出所有涉及同学的花名：张中、太奇、鸣世、曦轩、鸣珂、民瞻、九翎、渊虹、国相、晨炫、木晗、毅山、宋智、澄苍，是大家的合作与努力才让此书得以成行。同时感谢刘皎等编辑的辛苦工作，是你们把这本书呈献给读者，感谢你们！

同样要感谢支持我们工作的同事们：冰夷、王贲，有你们的帮助和支持才让我们有决心和毅力来完成这项工作。



目 录

第 1 章 Storm 基础	1
1.1 Storm 能做什么	2
1.2 Storm 特性	3
1.3 其他流计算系统	8
1.4 应用模式	13
第 2 章 Storm 初体验	17
2.1 本地环境搭建	18
2.2 Storm 集群	25

第 3 章 构建 Topology	41
3.1 Storm 基本概念	42
3.2 构建 Topology	53
3.3 小结	61
第 4 章 Topology 的并行度	62
4.1 并行元素	63
4.2 配置并行度	65
4.3 一个运行中 Topology 的例子	68
4.4 如何更新运行中的 Topology 的并行度	71
第 5 章 消息的可靠处理	73
5.1 简介	74
5.2 理解消息被完整处理	74
5.3 消息的生命周期	76
5.4 可靠相关的 API	79
5.5 高效地实现 tuple tree	84

5.6	选择合适的可靠性级别	87
5.7	集群的各级容错	89
5.8	小结	91
第 6 章 一致性事务		92
<hr/>		
6.1	简单设计一：强顺序流	93
6.2	简单设计二：强顺序 batch 流	95
6.3	CoordinateBolt 的原理	96
6.4	Transactional Topology	98
第 7 章 DRPC		105
<hr/>		
7.1	Storm DRPC	106
7.2	总体概述	106
7.3	LinearDRPCTopologyBuilder	108
7.4	本地模式 DRPC	110
7.5	远程模式 DRPC	111
7.6	一个复杂的例子	113

7.7	非线性 DRPC 拓扑	117
7.8	LinearDRPCTopologyBuilder 工作过程	117
7.9	高级进阶	118
第 8 章 Trident 的特性		119
8.1	理解 Trident	120
8.2	结合多个 Trident 任务	124
8.3	消费和生产 Field	126
8.4	State（状态保存）	128
8.5	Trident Topology 的执行过程	136
8.6	总结	137
第 9 章 Storm 实例		138
9.1	一个简单的实例	139
9.2	复杂一点的实例	150
9.3	其他	161

第 10 章 常见应用问题分析	162
10.1 性能问题排查与定位	163
10.2 系统中常见的问题与排查	167
10.3 业务问题的定位与排查	170

第 1 章

Storm 基础

1.1 Storm 能做什么

在大数据处理方面，相信大家已经对 Hadoop 耳熟能详了，Hadoop 处理的是存放在其分布式文件系统 HDFS 上的数据，Hadoop 使用磁盘作为中间交换的介质，在对海量数据进行离线分析时得心应手，但处理实时数据流却是力有未逮。

Storm 是一个开源的分布式实时计算系统，可以简单、可靠地处理大量的数据流。Storm 有很多使用场景，如实时分析、在线机器学习、持续计算、分布式 RPC、ETL，等等。Storm 支持水平扩展，具有高容错性，保证每个消息都会得到处理，而且处理速度很快（在一个小集群中，每个节点每秒可以处理数以百万计的消息）。Storm 的部署和运维都很便捷，而且更为重要的是可以使用任意编程语言来开发应用。

1.2 Storm 特性

Storm 有如下特点。

1. 编程模型简单

基于 Google Map/Reduce 来实现的 Hadoop 为开发者提供了 map、reduce 原语，使并行批处理程序变得非常简单和优美。同样，Storm 也为大数据的实时计算提供了一些简单优美的原语，这大大降低了开发并行实时处理任务的复杂性，帮助你快速、高效的开发应用。

2. 可扩展

在 Storm 集群中真正运行 Topology 的主要有三个实体：工作进程、线程和任务。Storm 集群中的每台机器上都可以运行多个工作进程，每个工作进程又可创建多个线程，每个线程可以执行多个任务，任务是真正进行数据处理的实体，Spout、Bolt 被开发出来就是作为一个或者多个任务的方式执行的。

因此，计算任务在多个线程、进程和服务器之间并行进行，支持灵活的水平扩展。

3. 高可靠性

Storm 可以保证 Spout 发出的每条消息都能被“完全处理”，这也是直接区别于其他实时系统的地方，如 Yahoo! S4。

请注意，Spout 发出的消息后续可能会触发产生成千上万条消息，可以形象地理解为一棵消息树，其中 Spout 发出的消息为树根，Storm 会跟踪这棵消息树的处理情况，只有当这棵消息树中的所有消息都被处理了，Storm 才会认为 Spout 发出的这个消息已经被“完全处理”。如果这棵消息树中的任何一个消息处理失败了，或者整棵消息树在限定的时间内没有“完全处理”，那么 Spout 发出的消息就会重发。

考虑到尽可能减少对内存的消耗，Storm 并不会跟踪消息树中的每个消息，而是采用了一些特殊的策略，它把消息树当作一个整体来跟踪，对消息树中所有消息的唯一 id 进行异或计算，通过是否为零来判定 Spout 发出的消息是否被“完全处理”，这极大地节约了内存和简化了判定逻辑，后面会对

这种机制进行详细介绍。

这种模式，每发送一个消息，都会同步发送一个 `ack` 或 `fail`，对于网络的带宽会有一些消耗，如果对于可靠性要求不高，可通过使用不同的 `emit` 接口关闭该模式。

上面所说的，Storm 保证了每个消息至少被处理一次，但是对于有些计算场合，会严格要求每个消息只被处理一次，幸而 Storm 的 0.7.0 版引入了事务性拓扑，解决了这个问题，本书后面会有详述。

4. 高容错性

如果在消息处理过程中出现了一些异常，Storm 会重新部署这个出问题的处理单元。Storm 保证一个处理单元永远运行（除非你显式的结束这个处理单元）。

当然，如果处理单元中存储了中间状态，那么当处理单元重新被 Storm 启动时，需要应用自己处理中间状态的恢复。