

新一代分布式的复杂事件处理系统解决方案

陈心咏

(福建富士通信息软件有限公司 福建 福州 350003)

【摘要】本文首先对复杂事件处理系统建设的驱动力进行介绍,然后结合系统的特性要求提出一套针对性的实现方案。该方案通过 Storm 结合策略处理引擎、消息队列和分布式缓存构建一套高可用性通用复杂事件处理系统,满足用户对复杂海量实时信息处理分析的需求。

【关键字】复杂事件处理;Storm;策略引擎;消息队列;分布式缓存;分布式数据库

一、引言

人类所在的世界,随时随地都在发生着大事小情,有些事情是无关紧要的,有些是和人们的切身安全和利益相关的,我们需要及时观察,及时处理响应周围时刻发生的事情。对于信息化系统也存在同样的问题,移动互联网时代正处于信息爆炸的时代,无时无刻不在发生着海量事件,传统的信息处理技术是一个低速复杂的事情,事情处理完后再对数据进行分析处理,然后从分析结果再做出决策,反馈执行新的决策动作,这时候这个最终决策动作已经完全滞后,其时效性大打折扣,甚至可能存在错误,这不是我们所期望的。

复杂事件处理是完全建立在实时处理流程上的,它基于预先总结配置的处理规则,并把规则配置在策略引擎中,对系统中的海量实时数据信息进行分析处理,并立即把处理结果和决策动作反馈在实时处理流程中,它是基于内存和消息队列的,是高速毫秒级的处理技术,这是移动互联网时代中系统架构的新的形式,对企业的快速反应、风险控制、实时处理具有深远的意义。

下面我们将介绍一套设计方案,主要通过利用开源框架 STORM 作为实时流式处理的主体框架,结合策略引擎、消息队列、分布式缓存技术能快速搭建一套高可用的分布式复杂事件处理系统。

二、复杂事件处理系统的要求和目标

- * 减少应用存储数据(在分析数据之前)造成的时延
- * 能够持续实时分析多个数据流,能够关联不同数据流中的事件进行处理
- * 能够迅速响应新发现的机会或危险
- * 能够利用已有的应用开发能力迅速开放新的高

性能、搞扩展的应用

- * 具备高可用性,在一台服务器发生危险的情况下,其它服务器能持续工作,确保应用系统处理的连贯性和实时性
- * 使用安全高效的内存管理技术,保证信息的可靠性和安全性

三、复杂事件处理系统的设计原理和实现方案

1、概念说明

在系统里,事件是指定义明确、分类清晰的业务信息的产生,每个事件的发生都带来一系列的信息,信息数据的变化和产生就是事件。因此在信息系统中它可以是对象的状态属性或者对象间动作触发。事件之间存在着时间先后关系、事件之间可以进行归类聚合、事件还可以被拆分、事件之间也会有依存关系和因果关系等。

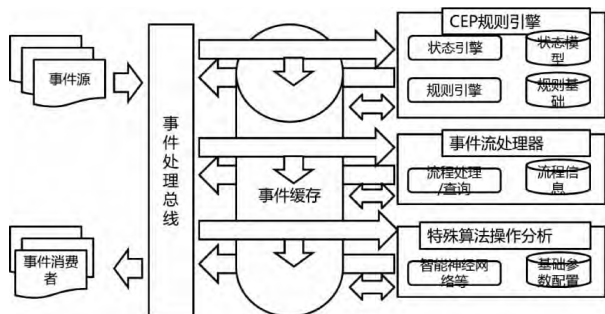
复杂事件处理系统是一种新兴的基于事件流的技术,它将系统数据看作不同类型的事件,通过策略引擎分析事件间的关系,建立不同的事件关系序列库,利用过滤、关联、聚合等技术,最终由简单无价值的事件(信息)产生高级事件或有价值事件,具备分析高速数据流和鉴别事件处理的能力。

2、系统功能框架

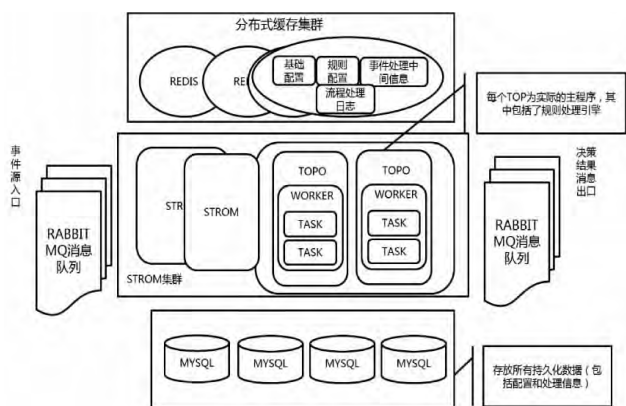
下图为系统的功能框架图。

实时数据作为事件源接入事件处理总线,然后存放在事件缓存中,事件流处理器主要针对事件流程中的原始信息进行记录并进行信息强化,以便策略引擎的二次分析,同时提供流程查询功能,策略引擎通过指定的规则,实现对消息数据处理,得到我们真正需要的有价值的信息,并通过这些有价值的信息,进行逻辑判断和分析决策,最终决定可执行的动作,特殊算法操作分析,主要针对一些非通用规则,既策略引

擎无法描述并配置的场景进行的特殊算法定制处理；最后把处理结果或者决策执行的动作通过事件处理总线返回给事件消费者。



3、系统技术架构



* 所有的事件源接入都放在 RABBITMQ 中

* STORM 集群通过消息队列获取数据进行后续实时消息处理。主要的处理过程都在 STORM 中,里面包括了策略引擎的处理内容

* 分布式缓存存放了基础配置、规则配置、事件处理中间信息、流程处理日志等各种实时处理信息,确保数据能够被高效的读取和应用

* MYSQL 数据库集群负责存放所有的持久化数据信息,以确保数据的安全性和可靠性。

* 所有的消息处理结果和决策动作结果会及时向末端的消息队列发送,同时,数据的处理决策结果也会存放一份在数据库中,以确保数据的安全性和可靠性。

四、复杂事件处理系统的关键技术

1、分布式实时消息流处理框架

Storm 是一个开源的分布式消息流处理框架,它有两种节点,一种是控制节点(Nimbus 节点),控制节点上面运行一个后台程序 Nimbus,它的作用类似 Hadoop 里面的 JobTracker。Nimbus 负责在集群里面分布代码,分配工作给机器,并且监控状态。另一种是工作节点(Supervisor 节点)。所有 Topology 任务的提交

必须在 Storm 客户端节点上进行,由 Nimbus 节点分配给其他 Supervisor 节点进行处理。Nimbus 节点首先将提交的 Topology 进行分片,分成一个个的 Task,并将 Task 和 Supervisor 相关的信息提交到集群上,Supervisor 会去集群上认领自己的 Task,通知自己的 Worker 进程进行 Task 的处理。

运行中的 Topology 主要由以下三个组件组成的: Worker processes (进程) Executors (threads)(线程), Tasks。每个运行的 Task 都由 Spout 和 Bolt 组成,这些 Spout 和 Bolt 是线程级的,实际上都在 Worker 进程中管理。系统的主要逻辑基于 Storm 建立,事件源的消息将通过分布于各个节点 Spout 读入,然后分发到多个节点的 Bolt 中进行分析处理,Bolt 中就是系统核心的事件策略处理引擎,因此每时每刻消息流入均能确保快速高效的并行处理分析,最终的决策动作,也将由末端的 Bolt 根据配置执行,然后把信息落到存储上

2、策略引擎模型

策略引擎模型包括事件、策略、段落、段落动作、条件、运算符、参考值、参考对象等实体,以下简要说明模型的关键实体:事件作为策略引擎处理的触发源,根据事件上的事件类型过滤匹配的处理策略,一个处理策略下挂多个段落条件,所有的事件原始信息通过预处理后,都将被引入段落条件中进行规则判断,当动作对应的段落树满足条件时,执行对应的决策动作。

策略引擎的主要处理逻辑如下:



* 获取一个批次的源事件信息。

* 对相同对象的事件进行合并处理(例如如果这一批事件中存在同一个帐户或用户的事件,则可以进行分析合并),判断事件的互斥关系,先后关系等

* 根据事件中的事件类型匹配对应策略。

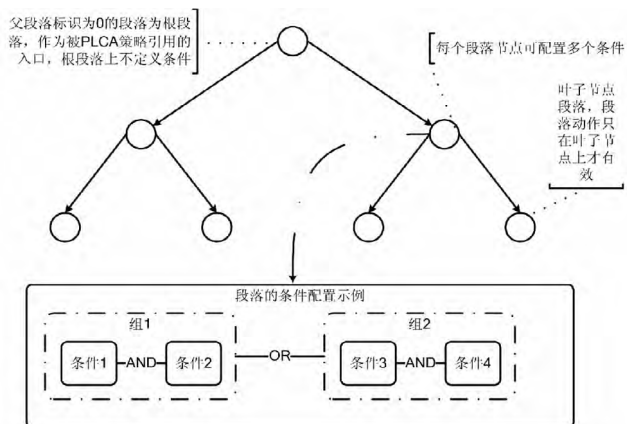
* 对策略下的条件段落树进行判断

* 条件中的参考信息可以引入各种可计算的参考对象,允许配置各类四则运算表达式和逻辑运算参考值

* 如果条件段落判断结果为真,并且段落下挂对应动作,则执行动作处理,否则进入下一个条件段落判断。

策略规则判断模型数据结构介绍如下:

* 以上为示意图,理论上段落可以有 N 个层级;



* 可以用数据结构中的树理解策略下挂的段落结构, 形成深度优先树形判断规则, 形成逻辑判断语句

* 同级的段落节点通过段落优先级字段确定执行顺序, 优先级越小越先执行;

* 同级的段落节点中, 按照优先级顺序执行所有段落节点;

* 在每个段落节点上可以配置多个条件, 且必须使用分组字段对条件进行分组, 同一组内的条件是 AND 关系, 条件组间是 OR 关系。

通过上述模型表达, 形成一套高效的策略规则处理引擎, 能对有效信息进行分析和判断, 并依据信息的含义和内容选择后续的处理动作。

3、高速的分布式缓存

基于 REDIS 的高速分布式缓存, 是流程中所有配置信息和处理过程信息的主要存放区域, 它是基于 KEY-VALUE 的方式存放, 能够达到对数据的快速访问, 同时自身的集群功能保证了数据的高可用性, 是实时消息系统中的最重要的数据存放区域。

4、消息中间件

消息中间件 RabbitMq 提供异步的持久和非持久消息的发布和订阅, 持久消息理论上应当具备绝对的可靠性, 通过消息中间件使得事件源端和处理事件末端与应用系统解耦, 提高整个流程的响应速度。

5、分布式数据库

分布式数据层采用 Mysql, 主要用作消息数据的持久化。它提供分库分表、读写分离、Sql 监控以及容灾容错的功能。这个组件不仅仅可以解决数据的拆分、读写分离问题, 还可以解决数据的多写, 多个 Slave 读取的负载均衡及容错。

五、系统特点总结

本文向大家介绍了一种通过开源技术框架结合自主研发的策略引擎建立一套高效分布式复杂事件处理系统的方案, 该方案与目前商用的系统相比部分指标可达到或接近商用系统, 大大降低了研发成本, 可用于支撑各类中小型实时消息类系统运营。

《福建电脑》投稿须知

《福建电脑》(国际标准刊号:ISSN 1673-2782, 国内统一刊号 CN 35-1115/TP)是福建省计算机学会主办、福建省科技厅主管的计算机学科综合月刊。欢迎投稿:

1. 论文所涉及的项目如国家或地方基金课题,请在来稿中注明课题项目、编号、来源。署作者真实姓名、工作单位、电话、通信地址、邮政编码和电子信箱。
2. 本刊设基金项目论文、研究与探讨、应用与开发、教学与管理等栏目。
3. 文稿要求论点明确、数据可靠、程序无误、文字准确简练。文稿一般不超过 5 千字(包括图、表、原程序、参考文献),并附中文摘要(200 字以内),中文题名一般不超过 20 个字。
4. 文稿要求层次分明、条理清晰。文章层次编号(二、三级标题)采用以下格式:0,1,2,...;1.1,1.2,1.2.1,...。四级标题以下不再设小标题。
5. 文稿中图表和照片一般不超过 4 个(包括应用程序框图等),文中图表及程序流程图或插图置于文内段落处,图表随文走,标明图位、中文图序和图题。文中附图宜用 WORD 制作,图幅宽一般不超过 8 cm,图中的文字采用 6 号字。
6. 本刊只受理电子邮件投稿,并在主题栏注明“新投稿”字样,如有修改稿请在主题栏写上稿件编号。文档请用 WORD 编排,并以附件形式发送。一封邮件控制在 2M 内,同一篇稿件请不要反复传送。
7. 编辑部收到作者稿件后,5 个工作日内反馈初审意见(电子邮件)。
8. 编辑部投稿邮箱:fjdnjb@vip.163.com 联系电话:0591-87814718 工作 QQ:2661504015
9. 为适应我国信息化建设,扩大本刊及作者知识信息交流渠道,本刊已被 CNKI 中国期刊全文数据库、万方数据库、维普资讯网收录,其作者文章著作权使用费与本刊稿酬一次性给付(已在收取发表费时折减和换算为杂志赠阅)。