



## (12) 发明专利申请

(10) 申请公布号 CN 104408656 A

(43) 申请公布日 2015. 03. 11

(21) 申请号 201410594354. 3

(22) 申请日 2014. 10. 29

(71) 申请人 中国建设银行股份有限公司

地址 100032 北京市西城区金融大街 25 号

(72) 发明人 王东

(74) 专利代理机构 广州三环专利代理有限公司

44202

代理人 郝传鑫 熊永强

(51) Int. Cl.

G06Q 40/02(2012. 01)

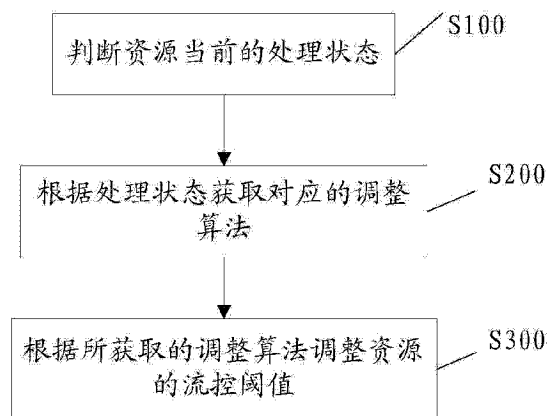
权利要求书2页 说明书6页 附图4页

### (54) 发明名称

动态调整流控阈值的方法及系统

### (57) 摘要

本发明提供一种动态调整流控阈值的方法及系统,其中,所述方法包括:判断资源当前的处理状态;根据所述处理状态获取对应的调整算法;根据所获取的调整算法调整所述资源的流控阈值。通过采用本发明所提供的方法,能够实现根据交易情况动态调整流控阈值,既实现基本流控功能,又最大限度地利用资源。



1. 一种动态调整流控阈值的方法,其特征在于,所述方法包括:  
判断资源当前的处理状态;  
根据所述处理状态获取对应的调整算法;  
根据所获取的调整算法调整所述资源的流控阈值。
2. 如权利要求 1 所述的方法,其特征在于,所述判断资源当前的处理状态包括:  
获取所述资源当前的交易总数和故障交易数;  
根据所述交易总数和所述故障交易数计算出故障率;  
将所述故障率与故障率标准进行比较,当所述故障率大于或等于所述故障率标准时,判断所述处理状态为故障状态,当所述故障率小于所述故障率标准时,判断所述处理状态为非故障状态。
3. 如权利要求 2 所述的方法,其特征在于,根据所述处理状态获取对应的调整算法包括:  
当判断所述处理状态为故障状态时,获取收缩算法;  
当判断所述处理状态为非故障状态时,获取恢复算法。
4. 如权利要求 3 所述的方法,其特征在于,根据收缩算法调整所述资源的流控阈值包括:  
获取收缩参数,其中所述收缩参数包括当前阈值、最小阈值、判定区间以及收缩率;  
将所述当前阈值与所述判定区间进行比较;  
根据所述比较的结果,将所述当前阈值设为相应的收缩目标值,其中若所述判定区间包含所述当前阈值,则所述收缩目标值为所述最小阈值,若所述判定区间不包含所述当前阈值,则所述收缩目标值为当前阈值 $\times(1-\text{收缩率})$ 。
5. 如权利要求 3 所述的方法,其特征在于,根据恢复算法调整所述资源的流控阈值包括:  
获取恢复参数,其中所述恢复参数包括当前阈值、最小阈值、最大阈值、初始阈值以及恢复率;  
将所述当前阈值与所述最小阈值进行比较;  
根据所述比较的结果,将所述当前阈值设为相应的恢复目标值,其中若所述当前阈值等于所述最小阈值,则所述恢复目标值为所述初始阈值与当前阈值 $\times(1+\text{恢复率})$ 中的较大值,若所述当前阈值不等于所述最小阈值,则所述恢复目标值为当前阈值 $\times(1+\text{恢复率})$ 与所述最大阈值中的较小值。
6. 如权利要求 1~5 中任一项所述的方法,其特征在于,  
按照预定的时间间隔执行所述判断资源当前的处理状态。
7. 一种动态调整流控阈值的系统,其特征在于,所述系统包括:  
判断模块,用于判断资源当前的处理状态;  
获取模块,用于根据所述判断模块判断的处理状态获取对应的调整算法;  
调整模块,用于根据所述获取模块所获取的调整算法调整所述资源的流控阈值。
8. 如权利要求 7 所述的系统,其特征在于,所述判断模块包括:  
获取单元,用于获取所述资源当前的交易总数和故障交易数;  
计算单元,用于根据所述获取单元获取的所述交易总数和所述故障交易数计算出故障

率；

判断单元,将所述计算单元计算出的故障率与故障率标准进行比较,当所述故障率大于或等于所述故障率标准时,判断所述处理状态为故障状态,当所述故障率小于所述故障率标准时,判断所述处理状态为非故障状态。

9. 如权利要求 8 所述的系统,其特征在于,所述获取模块包括：

收缩算法单元,用于当判断所述处理状态为故障状态时,获取收缩算法；

恢复算法单元,用于当判断所述处理状态为非故障状态时,获取恢复算法。

10. 如权利要求 9 所述的系统,其特征在于,所述调整模块包括：

第一获取组件,用于获取收缩参数,其中所述收缩参数包括当前阈值、最小阈值、判定区间以及收缩率；

第一比较组件,用于将所述第一获取组件获取的所述当前阈值与所述判定区间进行比较；

第一设置组件,用于根据所述第一比较组件所得出的所述比较的结果,将所述当前阈值设为相应的收缩目标值,其中若所述判定区间包含所述当前阈值,则所述收缩目标值为所述最小阈值,若所述判定区间不包含所述当前阈值,则所述收缩目标值为当前阈值 $\times(1 - \text{收缩率})$ 。

11. 如权利要求 9 所述的系统,其特征在于,所述调整模块还包括：

第二获取组件,用于获取恢复参数,其中所述恢复参数包括当前阈值、最小阈值、最大阈值、初始阈值以及恢复率；

第二比较组件,用于将所述第二获取组件所获取的所述当前阈值与所述最小阈值进行比较；

第二设置组件,用于根据所述第二比较组件所得出的所述比较的结果,将所述当前阈值设为相应的恢复目标值,其中若所述当前阈值等于所述最小阈值,则所述恢复目标值为所述初始阈值与当前阈值 $\times(1 + \text{恢复率})$ 中的较大值,若所述当前阈值不等于所述最小阈值,则所述恢复目标值为当前阈值 $\times(1 + \text{恢复率})$ 与所述最大阈值中的较小值。

12. 如权利要求 7 ~ 11 中任一项所述的系统,其特征在于,所述系统还包括：

定时模块,用于按照预定的时间间隔触发执行所述判断模块。

## 动态调整流控阈值的方法及系统

### 技术领域

[0001] 本发明涉及通信领域,更为具体而言,涉及动态调整流控阈值的方法及系统。

### 背景技术

[0002] 随着银行业对服务系统流量控制的不断完善与发展,目前,已经实现将某个服务系统的故障限制在一定范围内,不会影响其处理其他服务系统的交易。但是,由于流量控制是针对服务系统的交易并发量进行控制,而且目前服务系统交易并发量的阈值设置比较大,大于一些渠道系统的通讯接入连接数,那么当某个服务系统的交易出现大量超时,会导致核心平台的大量交易处理线程处于阻塞等待状态。虽然该服务系统的交易并发量并不一定达到了并发量阈值,但是渠道系统的通讯接入数可能已经全部用完,从而导致该渠道系统的所有后续交易都无法接入。

### 发明内容

[0003] 为解决上述技术问题,本发明提供了一种动态调整流控阈值的方法及系统,能够实现根据交易情况动态调整流控阈值,既实现基本流控功能,又最大限度地利用资源。

[0004] 根据本发明实施例的第一方面,提供了一种动态调整流控阈值的方法,所述方法包括:

[0005] 判断资源当前的处理状态;

[0006] 根据所述处理状态获取对应的调整算法;

[0007] 根据所获取的调整算法调整所述资源的流控阈值。

[0008] 根据本发明实施例的第二方面,提供了一种动态调整流控阈值的系统,所述系统包括:

[0009] 判断模块,用于判断资源当前的处理状态;

[0010] 获取模块,用于根据所述判断模块判断的处理状态获取对应的调整算法;

[0011] 调整模块,用于根据所述获取模块所获取的调整算法调整所述资源的流控阈值。

[0012] 实施本发明提供的动态调整流控阈值的方法及系统,能够实现根据交易情况动态调整流控阈值,既实现基本流控功能,又最大限度地利用资源。

### 附图说明

[0013] 图1是根据本发明实施方式的动态调整流控阈值的方法的流程图;

[0014] 图2示出图1所示的步骤S100的一种实施方式;

[0015] 图3示出图2所示的步骤S200的一种实施方式;

[0016] 图4是根据本发明实施方式的动态调整流控阈值的方法的另一流程图;

[0017] 图5是根据本发明实施方式的动态调整流控阈值的系统的结构示意图;

[0018] 图6示出了图5所示的判断模块100的一种实施方式的结构示意图;

[0019] 图7示出了图5所示的获取模块200的一种实施方式的结构示意图;

[0020] 图 8 示出了图 5 所示的调整模块 300 的一种实施方式的结构示意图。

### 具体实施方式

[0021] 为使本发明的实施例的目的、技术方案和优点更加清楚,下面将结合附图对本发明作进一步地详细描述。

[0022] 图 1 是根据本发明实施方式的动态调整流控阈值的方法的流程图。参见图 1,所述方法包括:

[0023] S100:判断资源当前的处理状态。

[0024] 其中,所述资源包括服务系统、渠道、交易等。

[0025] S200:根据所述处理状态获取对应的调整算法。

[0026] S300:根据所获取的调整算法调整所述资源的流控阈值。

[0027] 其中,可以按照预定的时间间隔执行所述步骤 S100。

[0028] 以上,对本发明提供的方法的整体流程进行了说明,下面结合实施例对本发明的方法的详细流程进行说明。

#### [0029] 【实施例 1】

[0030] 首先对所涉及到的术语进行解释说明:交易:银行系统中为完成某项业务而做的操作。流量控制:简称流控,主要指对交易并发量进行控制,使其不超过流控阈值。流控阈值:主要是某类交易并发量的限制值。动态流量控制:简称动态流控,指的是在流量控制实施过程中流控阈值可以动态调整。交易采样:在一段时间内对交易进行统计分析,计算出交易失败率,平均响应时间等信息。交易采样率:指多长时间统计分析一次交易情况。交易失败率:指在一个采样周期内交易失败数占所有交易的百分比。交易响应时间:一笔交易从发送请求到接到应答返回所需要的时间。

[0031] 在实施例 1 中,动态调整流控阈值的方法包括上述图 1 所示的处理方法的所有步骤,其中,如图 2 所示,步骤 S100 具体包括下述步骤:

[0032] S110:获取所述资源当前的交易总数和故障交易数。

[0033] S120:根据所述交易总数和所述故障交易数计算出故障率。

[0034] S130:将所述故障率与故障率标准进行比较,当所述故障率大于或等于所述故障率标准时,判断所述处理状态为故障状态,当所述故障率小于所述故障率标准时,判断所述处理状态为非故障状态。

#### [0035] 【实施例 2】

[0036] 在实施例 1 中,动态调整流控阈值的方法包括上述图 2 所示的处理方法的所有步骤,其中,如图 3 所示,步骤 S200 具体包括下述步骤:

[0037] S210:当判断所述处理状态为故障状态时,获取收缩算法。

[0038] S220:当判断所述处理状态为非故障状态时,获取恢复算法。

[0039] 其中,根据收缩算法调整所述资源的流控阈值包括:

[0040] 获取收缩参数,其中所述收缩参数包括当前阈值、最小阈值、判定区间以及收缩率,其中,所述判断区间可以设定为 [ 最小阈值,最大阈值 \*10% ],最小阈值为大于等于 1 且小于等于最大阈值的整数,默认为 1;

[0041] 将所述当前阈值与所述判定区间进行比较;

[0042] 根据所述比较的结果,将所述当前阈值设为相应的收缩目标值,其中若所述判定区间包含所述当前阈值,则所述收缩目标值为所述最小阈值,若所述判定区间不包含所述当前阈值,则所述收缩目标值为当前阈值 $\times(1-\text{收缩率})$ 。

[0043] 同时,根据恢复算法调整所述资源的流控阈值包括:

[0044] 获取恢复参数,其中所述恢复参数包括当前阈值、最小阈值、最大阈值、初始阈值以及恢复率,其中,初始阈值为大于最小阈值且小于最大阈值的整数,默认为最大阈值 $\times 10\%$ ;

[0045] 将所述当前阈值与所述最小阈值进行比较;

[0046] 根据所述比较的结果,将所述当前阈值设为相应的恢复目标值,其中若所述当前阈值等于所述最小阈值,则所述恢复目标值为所述初始阈值与当前阈值 $\times(1+\text{恢复率})$ 中的较大值,若所述当前阈值不等于所述最小阈值,则所述恢复目标值为当前阈值 $\times(1+\text{恢复率})$ 与所述最大阈值中的较小值。

[0047] 【实施例 3】

[0048] 在实施例 3 中,给出具体例子对动态调整流控阈值的方法进行说明。

[0049] 如图 4 所示,动态阈值调整模块按照时间周期从资源信息管理模块中获取目标资源在该周期内的采样交易总数和采样故障交易数,并计算故障率。通过配置文件设定需要动态阈值调整的资源条目,若某资源条目未在该配置文件中设定,则不进行动态阈值调整,即采用固定阈值进行流控。每个资源的统计周期由配置文件设置。每个周期结束时将采样数据清零。采样统计的触发功能由定时器实现,并采用独立的线程池执行。动态阈值调整算法分为收缩算法和恢复算法。由配置文件来设置某资源的故障率标准,当采样故障率大于或等于故障率标准时,采用收缩算法;当故障率小于故障率标准时,采用恢复算法,其中故障率就是交易超时率。

[0050] 收缩算法:由配置文件来设置某资源的收缩率。如果当前阈值不在区间[最小阈值,最大阈值 $\times 10\%$ ]之内,则收缩目标值=当前阈值 $\times(1-\text{收缩率})$ 。如果当前阈值在区间[最小阈值,最大阈值 $\times 10\%$ ]之内,则直接收缩至最小阈值。某资源的最小阈值由配置文件设置,为大于等于 1 且小于等于最大阈值的整数,默认为 1。当阈值收缩到最小值时,则不再收缩。

[0051] 恢复算法:由配置文件来设置某资源的恢复率。如果当前阈值不等于最小阈值,则恢复目标值=当前阈值 $\times(1+\text{恢复率})$ 与最大阈值中的较小值。如果当前阈值等于最小阈值,则恢复目标值取初始阈值和当前阈值 $\times(1+\text{恢复率})$ 的最大值。某资源的初始阈值由配置文件来设置,为大于最小阈值且小于最大阈值的整数,默认为最大阈值 $\times 10\%$ 。

[0052] 由配置文件来设置某资源的交易耗时标准,默认为交易超时时间标准。当交易响应时间大于所设交易耗时标准则认为是故障交易。如果交易在核心平台实际处理结果为超时,该交易也为故障交易。

[0053] 由配置文件来设置资源的采样持续时间和间隔时间,采样持续时间默认为 60 秒(参数化,可配置),间隔时间默认为 30 秒(即每次持续采样 60 秒,休息 30 秒,以此反复)。

[0054] 上述过程的采样统计窗口采用时间窗口,在该时间窗口内统计目标资源的采样交易总数和采样故障交易数,并计算故障率(故障率=采样故障交易数/采样交易总数)。每个资源的采样统计窗口由配置文件设置,默认为 120 秒。每个采样统计窗口结束时将采样

数据清零,清零工作由动态阈值调整模块完成。

[0055] 由配置文件来设置某资源的故障率标准,默认值(参数化,可配置)为 20%。在每次采样统计计算之后,将故障率与故障率标准进行比较;当故障率大于或等于故障率标准时,进行阈值动态收缩;当故障率小于故障率标准时,进行阈值动态恢复。

[0056] 本发明支持等比收缩算法,即由配置文件来设置某资源的收缩率,默认(参数化,可配置)为 50%。阈值动态收缩时按比例收缩,收缩目标值=当前阈值\*(1-收缩率)。针对上述收缩算法,由配置文件来设置某资源的最小阈值,默认为 1。阈值动态收缩时不能小于最小阈值。最小阈值不小于 1,不大于流量控制最大阈值。针对上述收缩算法,当收缩时,依照收缩率按比例收缩。如果当前阈值在区间[最小阈值,最大阈值\*10%],直接收缩至最小阈值。

[0057] 本发明支持等比恢复算法,即由配置文件来设置某资源的恢复率,默认为 100%。阈值动态恢复时按比例恢复,恢复目标值=当前阈值\*(1+恢复率)。针对上述恢复算法,由配置文件来设置某资源的最低恢复比值,默认为 10%及最大阈值。最低恢复阈值=最低恢复比值\*最大阈值。阈值动态恢复目标值不能小于最低恢复阈值。如果按照上面算法得到的恢复目标值小于最低恢复阈值,则将最低恢复阈值作为恢复目标值。最低恢复阈值不大于最大阈值。

[0058] 阈值的动态调整应该与交易执行松耦合,采用独立线程,在每个采样周期结束时进行故障率计算和阈值动态调整。并且可以采用定时任务来执行阈值动态调整,每批相同采样周期为一个独立定时任务,该独立定时任务完成所有属于该采样周期目标资源的阈值动态调整工作并提供 INFO 级别日志信息。

[0059] 由配置文件来设置某资源的阈值动态调整开关,为必配项,on 表示允许动态调整,off 表示禁止动态调整。根据开关状态决定当前资源是否采集数据和动态调整。

[0060] 可以通过管理工具手动刷新配置文件和设置资源阈值。提供操作日志信息。

[0061] 本发明支持不同资源(包括服务系统、渠道、交易等)的并发量阈值调整,实现对分行、服务系统(或两者组合)以及节点号的并发量阈值动态调整。

[0062] 综上,当资源处理缓慢时,能够自动减少当前的流控阈值,以保证其他资源能够正常处理,以达到故障隔离的效果,其次当资源处理正常时,能自动恢复当前的流控阈值,直至恢复到最大阈值。

[0063] 本发明还可以根据资源的不同情况选择不同的线型控制策略:1、逐渐收缩和恢复;2、两级阶梯(高、低两种并发数);3、并发数不变化。每一服务系统资源选择其中一种策略进行流量控制。在使用上,交易量非常大且后台交易响应很快的服务系统(比如主机服务系统),采用逐渐收缩和恢复的控制策略;交易量较大且后台交易响应不是很快的服务系统(比如龙网服务系统),采用两级阶梯的控制策略;交易量很小的服务系统采用并发数不变化的控制策略。

[0064] 图 5 是根据本发明实施方式的动态调整流控阈值的系统的结构示意图。参见图 5,所述系统包括:

[0065] 判断模块 100,用于判断资源当前的处理状态。

[0066] 其中,所述资源包括服务系统、渠道、交易等。

[0067] 获取模块 200,用于根据所述判断模块 100 判断的处理状态获取对应的调整算法。

[0068] 调整模块 300, 用于根据所述获取模块 200 所获取的调整算法调整所述资源的流控阈值。

[0069] 其中, 所述系统还可以包括定时模块, 用于按照预定的时间间隔触发执行所述判断模块 100。

[0070] 如图 6 所示, 在本发明的实施方式中, 判断模块 100 可以包括:

[0071] 获取单元 110, 用于获取所述资源当前的交易总数和故障交易数。

[0072] 计算单元 120, 用于根据所述获取单元 110 获取的所述交易总数和所述故障交易数计算出故障率。

[0073] 判断单元 130, 将所述计算单元 120 计算出的故障率与故障率标准进行比较, 当所述故障率大于或等于所述故障率标准时, 判断所述处理状态为故障状态, 当所述故障率小于所述故障率标准时, 判断所述处理状态为非故障状态。

[0074] 如图 7 所示, 在本发明的实施方式中, 获取模块 200 可以包括:

[0075] 收缩算法单元 210, 用于当判断所述处理状态为故障状态时, 获取收缩算法;

[0076] 恢复算法单元 220, 用于当判断所述处理状态为非故障状态时, 获取恢复算法。

[0077] 如图 8 所示, 在本发明的实施方式中, 调整模块 300 可以包括:

[0078] 第一获取组件 310, 用于获取收缩参数, 其中所述收缩参数包括当前阈值、最小阈值、判定区间以及收缩率, 其中, 所述判定区间可以设定为 [ 最小阈值, 最大阈值 \* 10% ], 最小阈值为大于等于 1 且小于等于最大阈值的整数, 默认为 1;

[0079] 第一比较组件 320, 用于将所述第一获取组件 310 获取的所述当前阈值与所述判定区间进行比较;

[0080] 第一设置组件 330, 用于根据所述第一比较组件 320 所得出的所述比较的结果, 将所述当前阈值设为相应的收缩目标值, 其中若所述判定区间包含所述当前阈值, 则所述收缩目标值为所述最小阈值, 若所述判定区间不包含所述当前阈值, 则所述收缩目标值为当前阈值 \* (1 - 收缩率)。

[0081] 第二获取组件 340, 用于获取恢复参数, 其中所述恢复参数包括当前阈值、最小阈值、最大阈值、初始阈值以及恢复率, 其中, 初始阈值为大于最小阈值且小于最大阈值的整数, 默认为最大阈值 \* 10%;

[0082] 第二比较组件 350, 用于将所述第二获取组件 340 所获取的所述当前阈值与所述最小阈值进行比较;

[0083] 第二设置组件 360, 用于根据所述第二比较组件 350 所得出的所述比较的结果, 将所述当前阈值设为相应的恢复目标值, 其中若所述当前阈值等于所述最小阈值, 则所述恢复目标值为所述初始阈值与当前阈值 \* (1 + 恢复率) 中的较大值, 若所述当前阈值不等于所述最小阈值, 则所述恢复目标值为当前阈值 \* (1 + 恢复率) 与所述最大阈值中的较小值。

[0084] 实施本发明提供的动态调整流控阈值的方法及系统, 能够实现根据交易情况动态调整流控阈值, 既实现基本流控功能, 又最大限度地利用资源。

[0085] 通过以上的实施方式的描述, 本领域的技术人员可以清楚地了解到本发明可借助软件结合硬件平台的方式来实现, 当然也可以全部通过硬件来实施。基于这样的理解, 本发明的技术方案对背景技术做出贡献的全部或者部分可以以软件产品的形式体现出来, 该计算机软件产品可以存储在存储介质中, 如 ROM/RAM、磁碟、光盘等, 包括若干指令用以使得一



台计算机设备（可以是个人计算机，服务器，智能手机或者网络设备等）执行本发明各个实施例或者实施例的某些部分所述的方法。

[0086] 本发明说明书中使用的术语和措辞仅仅为了举例说明，并不意味构成限定。本领域技术人员应当理解，在不脱离所公开的实施方式的基本原理的前提下，对上述实施方式中的各细节可进行各种变化。因此，本发明的范围只由权利要求确定，在权利要求中，除非另有说明，所有的术语应按最宽泛合理的意思进行理解。

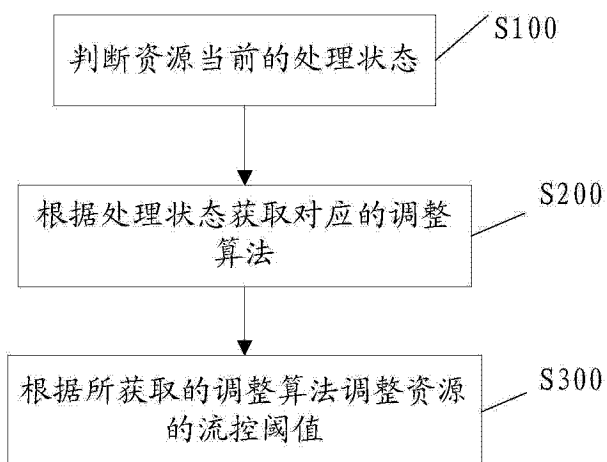


图 1

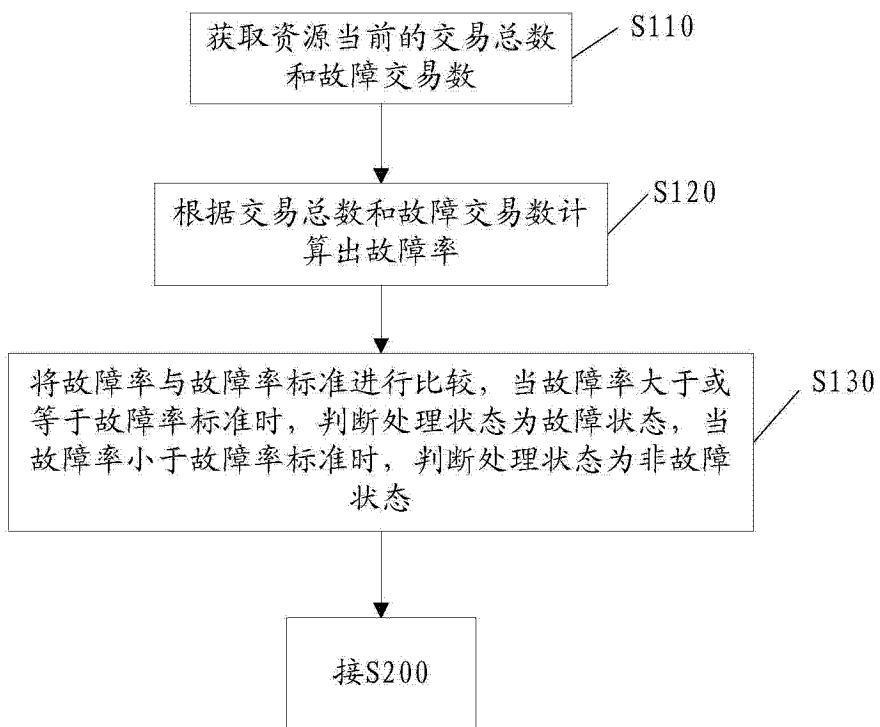


图 2

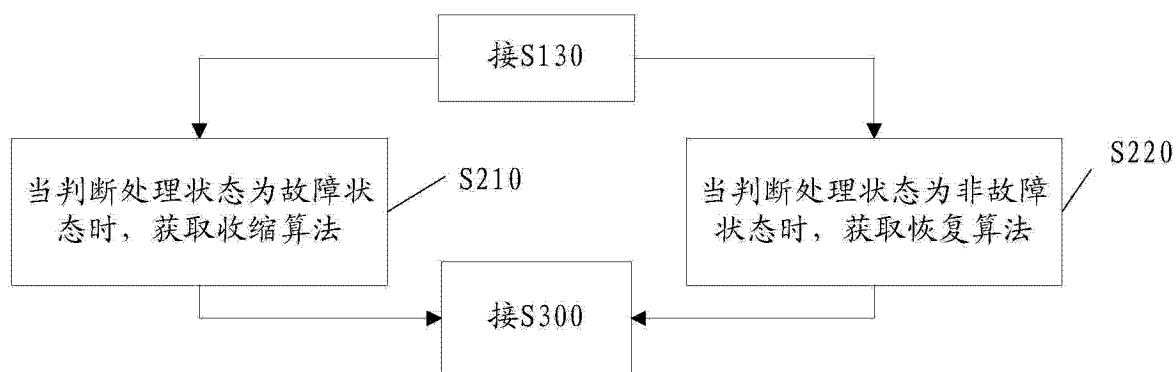


图 3

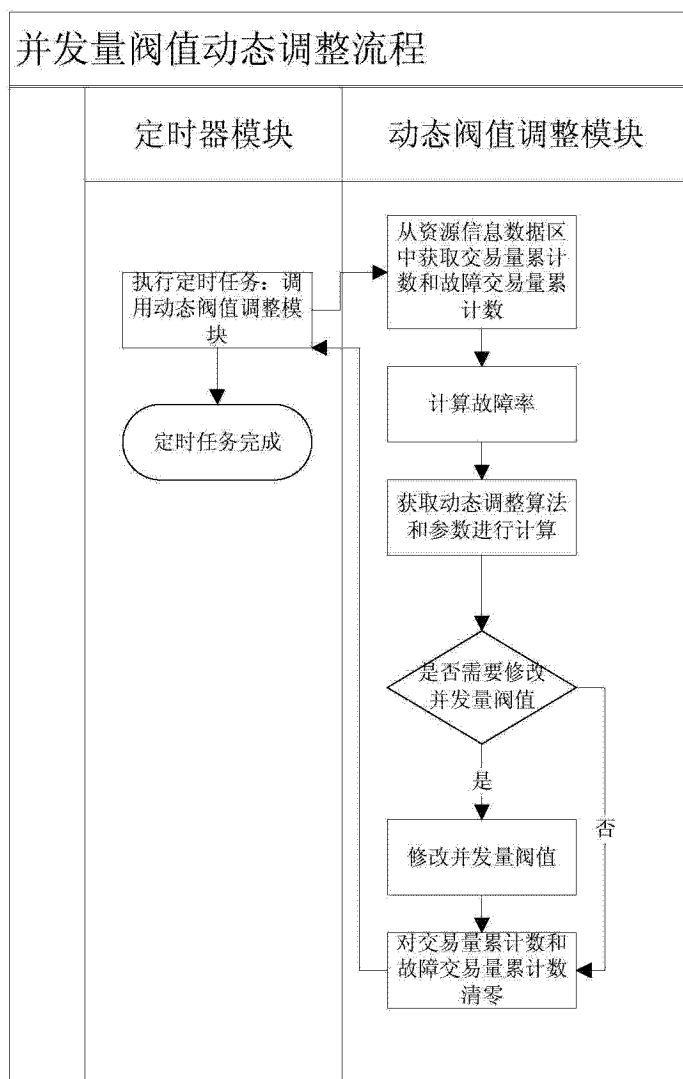


图 4

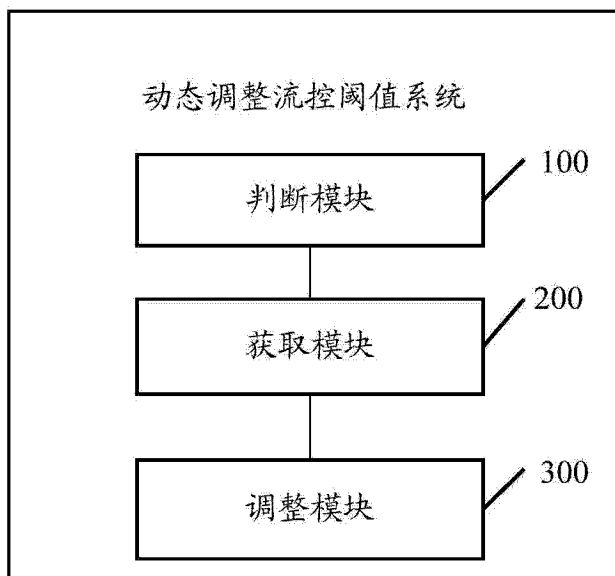


图 5

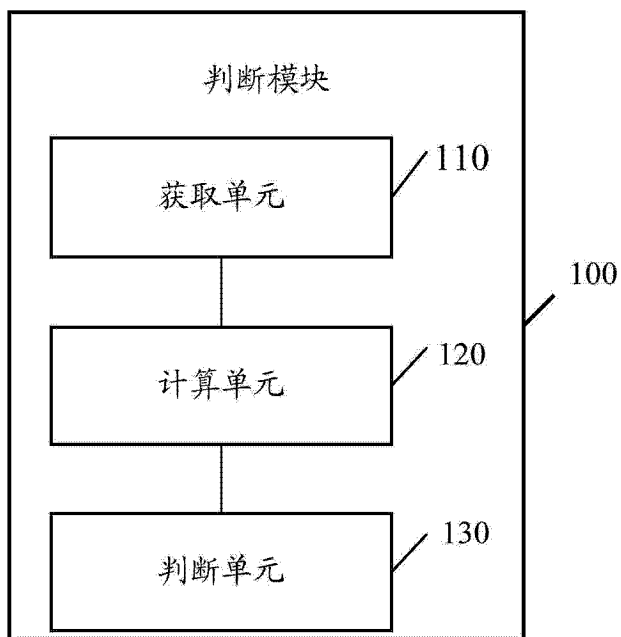


图 6

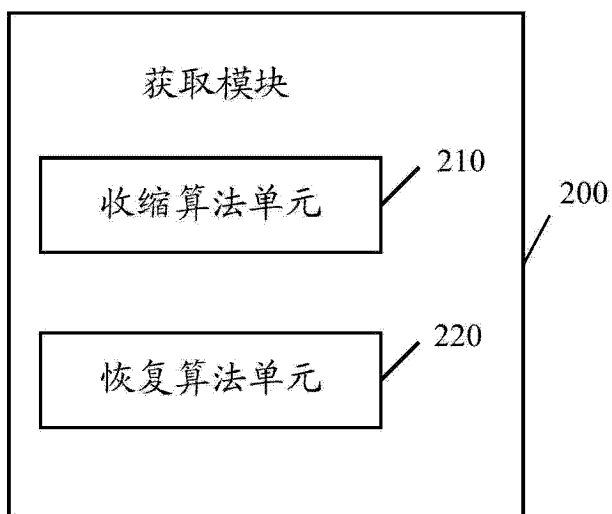


图 7

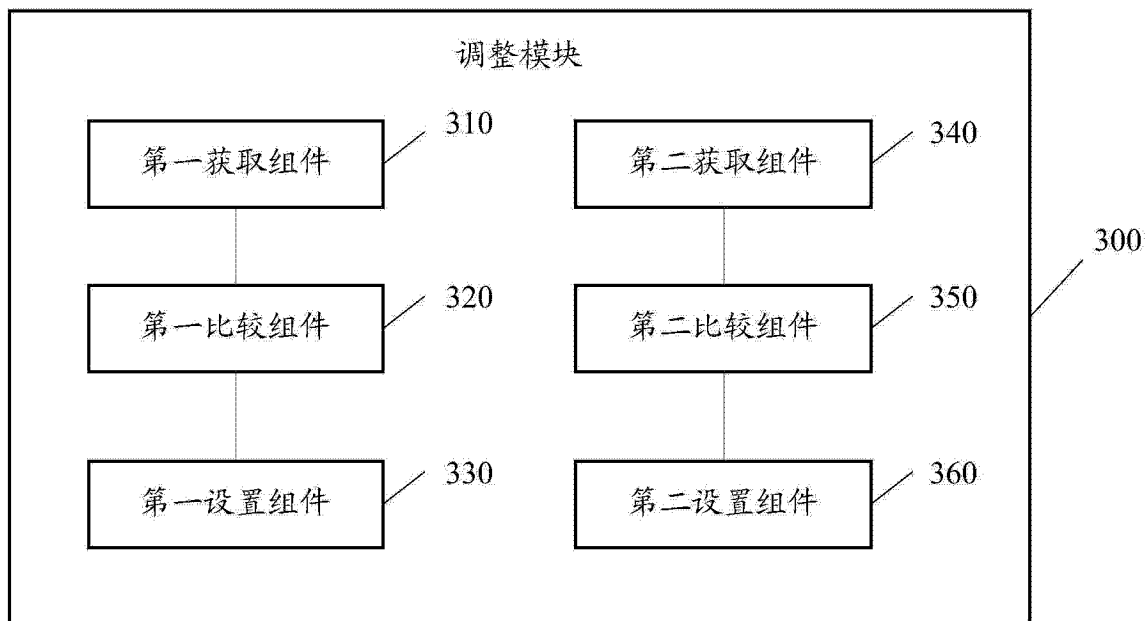


图 8