(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 112395155 B
(45)授权公告日 2024.11.01

(21)申请号 202011232331.X

(22)申请日 2020.11.06

(65)同一申请的已公布的文献号
申请公布号 CN 112395155 A

(43)申请公布日 2021.02.23

(73)专利权人 微民保险代理有限公司
地址 518063 广东省深圳市南山区粤海街道深南大道9996号松日鼎盛大厦25楼

(72)发明人 黄力平

(74)专利代理机构 深圳智汇远见知识产权代理有限公司 44481
专利代理师 李雪鹃 王旭

(51)Int.Cl.
G06F 11/30(2006.01)

(56)对比文件
CN 109062754 A,2018.12.21

审查员 李珍珍

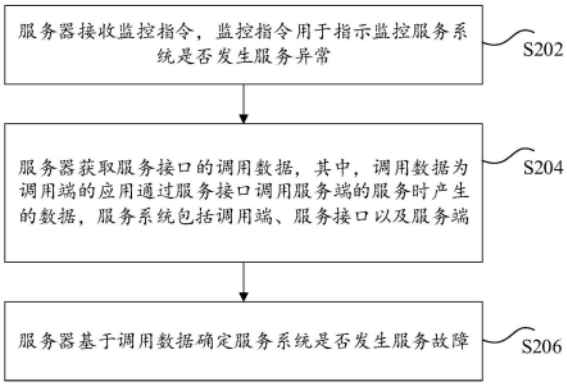
权利要求书3页 说明书12页 附图6页

(54)发明名称

服务的监控方法和装置、存储介质、电子装置

(57)摘要

本申请公开了一种服务的监控方法和装置、存储介质、电子装置。其中,该方法包括:接收监控指令,其中,所述监控指令用于指示监控服务系统是否发生服务异常;响应于所述监控指令,获取服务接口的调用数据,其中,所述调用数据为调用端的应用通过所述服务接口调用服务端的服务时产生的数据,所述服务系统包括所述调用端、所述服务接口以及所述服务端;基于所述调用数据确定所述服务系统是否发生服务故障。本申请解决了相关技术中的故障检测时效性较差的技术问题。



CN 112395155 B

1.一种服务的监控方法,其特征在于,包括:

接收监控指令,其中,所述监控指令用于指示监控服务系统是否发生服务异常;

响应于所述监控指令,获取多个服务接口的调用数据,其中,所述调用数据为调用端的应用通过所述服务接口调用服务端的服务时产生的数据,所述服务系统包括所述调用端、所述服务接口以及所述服务端;

基于所述调用数据确定所述服务系统是否发生服务故障;

其中,所述调用数据包括调用频率,多个所述服务接口包括当前处理的目标接口,服务故障包括位于所述目标接口的调用端的目标应用发生的故障,根据所述目标接口的调用频率确定所述目标应用是否发生故障包括以下之一:

通过比较所述目标接口的调用频率和频率阈值确定所述目标应用是否发生故障;

通过所述目标接口的调用频率在频率数据集中的位置确定所述目标应用是否发生故障,其中,所述频率数据集包括所述目标接口的调用频率;

通过比较第一子周期内的调用频率与第二子周期内的调用频率确定所述目标应用是否发生故障,其中,所述第一子周期为当前获取的调用频率所在的时间子周期,所述第二子周期为所述第一子周期的前一子周期;

通过比较所述第一子周期内的调用频率与第三子周期内的调用频率确定所述目标应用是否发生故障,其中,所述第一子周期在第一周期内的位置与所述第三子周期在第二周期内的位置相同,所述第二周期为所述第一周期的前一时间周期。

2.根据权利要求1所述的方法,其特征在于,基于所述调用数据确定所述服务系统是否发生服务故障包括:

根据多个所述服务接口中每个所述服务接口的调用数据确定所述服务接口、位于所述服务接口的调用端的应用以及位于所述服务接口的服务端的服务中的至少之一是否发生故障。

3.根据权利要求2所述的方法,其特征在于,根据所述服务接口的调用数据确定所述服务接口、位于所述服务接口的调用端的应用以及位于所述服务接口的服务端的服务中的至少之一是否发生故障包括:

在目标接口收到的请求数据发生异常的情况下,确定目标应用发生故障,其中,所述调用数据包括所述请求数据;

在所述目标接口收到的请求数据正常且与所述请求数据对应的响应数据发生异常的情况下,确定目标服务发生异常,其中,所述目标服务为位于所述目标接口的服务端的服务,所述调用数据包括所述响应数据;

在所述目标接口对调用数据的传输频率低于第一阈值且与提供所述目标服务的其他服务接口对调用数据的传输频率高于第二阈值的情况下,确定所述目标接口发生异常,其中,所述第二阈值高于所述第一阈值。

4.根据权利要求1所述的方法,其特征在于,所述频率阈值包括最低频率阈值和最高频率阈值,其中,在通过比较所述目标接口的调用频率和频率阈值确定所述目标应用是否发生故障之前,所述方法还包括:

统计所述服务系统正常运行时所述目标接口的最低调用频率和最高调用频率;

将所述最低调用频率与缩小系数之间的乘积作为所述最低频率阈值,将所述最高调用

频率与放大系数之间的乘积作为所述最高频率阈值。

5.根据权利要求1所述的方法,其特征在于,在通过所述目标接口的调用频率在频率数据集中的位置确定所述目标应用是否发生故障之前,所述方法还包括:

获取所述目标接口的历史调用频率的最小值和最大值;

根据所述最小值和所述最大值确定第一四分位数Q1、中位数Q2以及第三四分位数Q3,其中,所述第一四分位数为所述最大值和所述最小值的差值的四分之一与所述最小值之和,所述中位数为所述最大值与所述最小值的和的二分之一,所述第三四分位数为所述最大值和所述最小值的差值的四分之三与所述最小值之和;

获取所述第三四分位数与所述第一四分位数的差值为四分位全距IQR;

获取所述第一四分位数与所述四分位全距的X倍之间的差值为频率下限(Q1-X*IQR),并获取所述第三四分位数与所述四分位全距的Y倍之间的和为频率上限(Q4+Y*IQR),X和Y为调控参数。

6.根据权利要求1所述的方法,其特征在于,在基于所述调用数据确定所述服务系统是否发生服务故障之后,所述方法还包括:

将当前获取到的目标接口的调用数据保存至数据库中,其中,所述数据库按照配置的过期时间将过期的调用数据删除;和/或,在确定所述服务系统发生服务故障的情况下,推送报警信息。

7.一种服务的监控装置,其特征在于,包括:

接收单元,用于接收监控指令,其中,所述监控指令用于指示监控服务系统是否发生服务异常;

获取单元,用于响应于所述监控指令,获取多个服务接口的调用数据,其中,所述调用数据为调用端的应用通过所述服务接口调用服务端的服务时产生的数据,所述服务系统包括所述调用端、所述服务接口以及所述服务端;

监控单元,用于基于所述调用数据确定所述服务系统是否发生服务故障;

其中,所述调用数据包括调用频率,多个所述服务接口包括当前处理的目标接口,服务故障包括位于所述目标接口的调用端的目标应用发生的故障,所述监控单元还用于根据所述目标接口的调用频率确定所述目标应用是否发生故障包括以下之一:

通过比较所述目标接口的调用频率和频率阈值确定所述目标应用是否发生故障;

通过所述目标接口的调用频率在频率数据集中的位置确定所述目标应用是否发生故障,其中,所述频率数据集包括所述目标接口的调用频率;

通过比较第一子周期内的调用频率与第二子周期内的调用频率确定所述目标应用是否发生故障,其中,所述第一子周期为当前获取的调用频率所在的时间子周期,所述第二子周期为所述第一子周期的前一子周期;

通过比较所述第一子周期内的调用频率与第三子周期内的调用频率确定所述目标应用是否发生故障,其中,所述第一子周期在第一周期内的位置与所述第三子周期在第二周期内的位置相同,所述第二周期为所述第一周期的前一时间周期。

8.一种计算机可读的存储介质,其特征在于,所述存储介质包括存储的程序,其中,所述程序运行时执行上述权利要求1至6任一项中所述的方法。

9.一种电子装置,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运

行的计算机程序,其特征在于,所述处理器通过所述计算机程序执行上述权利要求1至6任一项中所述的方法。

10.一种电子装置,包括计算机程序/指令,其特征在于,所述计算机程序/指令被处理器执行时实现权利要求1至6任一项中所述的方法。

服务的监控方法和装置、存储介质、电子装置

技术领域

[0001] 本申请涉及互联网领域，具体而言，涉及一种服务的监控方法和装置、存储介质、电子装置。

背景技术

[0002] 在生产环境中，各种应用系统会被部署在众多的服务器或容器上，这些系统在运行过程中会输出各种日志，来反映系统状态、反馈业务执行情况等，通过采集和分析这些日志信息，可以进行故障分析与监控告警。

[0003] 相关技术中，需要把问题积累到一定程度以后，取出数据并进行分析才能识别出故障并发出告警，在时效性方面有一些滞后。

[0004] 针对上述的问题，目前尚未提出有效的解决方案。

发明内容

[0005] 本申请实施例提供了一种服务的监控方法和装置、存储介质、电子装置，以至少解决相关技术中的故障检测时效性较差的技术问题。

[0006] 根据本申请实施例的一个方面，提供了一种服务的监控方法，包括：接收监控指令，其中，监控指令用于指示监控服务系统是否发生服务异常；响应于监控指令，获取服务接口的调用数据，其中，调用数据为调用端的应用通过服务接口调用服务端的服务时产生的数据，服务系统包括调用端、服务接口以及服务端；基于调用数据确定服务系统是否发生服务故障。

[0007] 根据本申请实施例的另一方面，还提供了一种服务的监控装置，包括：接收单元，用于接收监控指令，其中，监控指令用于指示监控服务系统是否发生服务异常；获取单元，用于响应于监控指令，获取服务接口的调用数据，其中，调用数据为调用端的应用通过服务接口调用服务端的服务时产生的数据，服务系统包括调用端、服务接口以及服务端；监控单元，用于基于调用数据确定服务系统是否发生服务故障。

[0008] 根据本申请实施例的另一方面，还提供了一种存储介质，该存储介质包括存储的程序，程序运行时执行上述的方法。

[0009] 根据本申请实施例的另一方面，还提供了一种电子装置，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，处理器通过计算机程序执行上述的方法。

[0010] 在本申请实施例中，在需要监控服务系统是否发生服务异常时，可获取服务接口的调用数据，基于调用数据确定服务系统是否发生服务故障，由于获取的调用数据是实时数据，进而可以在第一时间确定是否发生故障，可以解决相关技术中的故障检测时效性较差的技术问题，进而达到提高检测时效性的技术效果。

附图说明

[0011] 此处所说明的附图用来提供对本申请的进一步理解,构成本申请的一部分,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。在附图中:

[0012] 图1是根据本申请实施例的服务的监控方法的硬件环境的示意图;

[0013] 图2是根据本申请实施例的一种可选的服务的监控方法的流程图;

[0014] 图3是根据本申请实施例的一种可选的服务的监控方案的示意图;

[0015] 图4是根据本申请实施例的一种可选的服务的监控方案的示意图;

[0016] 图5是根据本申请实施例的一种可选的服务的监控方案的示意图;

[0017] 图6是根据本申请实施例的一种可选的服务的监控方案的示意图;

[0018] 图7是根据本申请实施例的一种可选的指标频率的示意图;

[0019] 图8是根据本申请实施例的一种可选的指标的箱线图的示意图;

[0020] 图9是根据本申请实施例的一种可选的服务的监控装置的示意图;

[0021] 以及

[0022] 图10是根据本申请实施例的一种终端的结构框图。

具体实施方式

[0023] 为了使本技术领域的人员更好地理解本申请方案,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申请一部分的实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都应当属于本申请保护的范围。

[0024] 需要说明的是,本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本申请的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0025] 根据本申请实施例的一方面,提供了一种服务的监控方法的实施例。该方法可以应用在对业务进行实时监控告警的业务场景中,比如调用端、服务接口、服务端发生故障的情况下,可以采用短信和邮件的告警方式发送至代码工程师以提示发生故障,便于代码工程师准确定位问题,进行故障排除和修复。

[0026] 可选地,在本申请实施例中,上述方法可以应用于如图1所示的硬件环境中。如图1所示,终端102中可以包含有存储器104、处理器106和显示器108(可选部件)。终端102可以通过网络110与服务器112进行通信连接,该服务器112可用于为终端或终端上安装的客户端提供服务(如游戏服务、应用服务等),可在服务器112上或独立于服务器112设置数据库114,用于为服务器112提供数据存储服务。此外,服务器112中可以运行有处理引擎116,该处理引擎116可以用于执行由服务器112所执行的步骤。

[0027] 可选地,终端102可以但不限于为可以计算数据的终端,如移动终端(例如手机、平

板电脑)、笔记本电脑、PC(Personal Computer,个人计算机)机等终端上,上述网络可以包括但不限于无线网络或有线网络。其中,该无线网络包括:蓝牙、WIFI(Wireless Fidelity,无线保真)及其他实现无线通信的网络。上述有线网络可以包括但不限于:广域网、城域网、局域网。上述服务器112可以包括但不限于任何可以进行计算的硬件设备。

[0028] 此外,在本实施例中,上述方法还可以但不限于应用于处理能力较强大的独立的处理设备中,而无需进行数据交互。例如,该处理设备可以但不限于为处理能力较强大的终端设备,即,上述方法中的各个操作可以集成在一个独立的处理设备中。上述仅是一种示例,本实施例中对此不作任何限定。

[0029] 在本申请实施例中,上述方法可以由服务器112来执行,也可以是由服务器112和终端102共同执行,后续以由服务器112来执行为例进行说明。图2是根据本申请实施例的一种可选的服务的监控方法的流程图,如图2所示,该方法可以包括以下步骤:

[0030] 步骤S202,服务器接收监控指令,监控指令用于指示监控服务系统是否发生服务异常。

[0031] 步骤S204,响应于监控指令,服务器获取服务接口的调用数据,其中,调用数据为调用端的应用通过服务接口调用服务端的服务时产生的数据,服务系统包括调用端、服务接口以及服务端。

[0032] 调用端即调用服务的应用所在的一侧,该应用具体可以为用户在终端上使用的应用,也可以是调用基础服务的另一个服务;服务端即提供服务的一侧,可以部署在服务器或者云上。

[0033] 步骤S206,服务器基于调用数据确定服务系统是否发生服务故障。

[0034] 相关技术中,利用日志分析故障,需要进行问题累计,累计完成后在需要检测故障时,收集相对于时间而言具有一定信息延迟的日志,然后通过对日志的分析确定是否发生故障,其中的问题累计和日志分析需要消耗时间,且分析的日志所记载的信息具有一定的延时性,从而造成故障检测的时效性较差,而在本申请的技术方案中,通过上述步骤,在需要监控服务系统是否发生服务异常时,可获取服务接口的调用数据,基于调用数据确定服务系统是否发生服务故障,由于获取的调用数据是实时数据,进而可以在第一时间确定是否发生故障,可以解决相关技术中的故障检测时效性较差的技术问题,进而达到提高检测时效性的技术效果。下文结合具体的实现方案进一步详述本申请的技术方案:

[0035] 步骤11,在需要监控服务系统是否发生服务异常时,触发监控指令。

[0036] 例如,1)可以启动一个定时器,定时器记满一个时间周期(如1分钟、1小时等)后,触发监控指令;2)用户预先配置好脚本,在脚本中配置触发监控指令的条件,当满足条件时触发监控指令;3)用户在后台进行人工操作,根据需求触发监控指令。

[0037] 步骤12,服务器获取多个服务接口的调用数据。

[0038] 针对每个服务,可以经由一个或者多个服务接口接入,经由该接口路由的数据(如请求数据、响应数据)均可由监控侧获取。

[0039] 步骤13,针对每个服务接口,可根据服务接口的调用数据确定服务接口、位于服务接口的调用端的应用以及位于服务接口的服务端的服务中的至少之一是否发生故障。

[0040] 在该方案中,可对每个服务接口的调用数据进行如下处理:将产生当前待处理的调用数据的服务接口作为目标接口;通过逐个分析目标接口的调用数据确定目标接口、目

标应用以及目标服务中的至少之一是否发生故障,目标应用为位于目标接口的调用端的应用,目标服务为位于目标接口的服务端的服务。

[0041] 例如,若收到的请求数据发生异常(如请求数据量变少、请求数据的数据格式不正确等),则可以判定是调用端的应用发生故障;若收到的请求数据正常,但是响应数据发生异常(如响应数据量变少、响应数据的数据格式不正确等),则可以判定是服务端的服务发生故障;若一个服务接口几乎没有接收到请求数据和响应数据(即目标接口对调用数据的传输频率低于第一阈值),而提供相同服务的其他接口是正常的(即其他服务接口对调用数据的传输频率高于第二阈值,第一阈值为根据第二阈值确定的,如第一阈值为第二阈值的10%),则可以判定是该服务接口发生故障。后文以确认应用侧的应用是否发生故障为例进行说明:

[0042] 可选地,根据多个服务接口中每个服务接口的调用数据确定位于服务接口的调用端的应用是否发生故障时,调用数据包括调用频率,多个服务接口包括当前处理的目标接口(下文以单个目标接口为例描述了具体的故障确定方案,在实际处理时,若对各个服务接口采用串行处理的方式,则依次将多个服务接口作为目标接口并按照如下方式进行处理即可,在实际处理时,若对各个服务接口采用并行处理的方式,则依据并行量取多个服务接口中与并行量同等数量的服务接口作为目标接口并将每个目标接口分别按照如下方式进行处理即可),根据目标接口的调用频率确定目标应用是否发生故障包括以下几种方式:

[0043] 方式一,通过比较目标接口的调用频率和频率阈值确定目标应用是否发生故障。

[0044] 在通过比较目标接口的调用频率和频率阈值确定目标应用是否发生故障之前,预先统计服务系统正常运行时目标接口的最低调用频率和最高调用频率;考虑到,以上统计可能存在一定的片面性,可以将最低调用频率与缩小系数(如0.7)之间的乘积作为最低频率阈值,将最高调用频率与放大系数(如1.5)之间的乘积作为最高频率阈值。

[0045] 比较时,若目标接口的调用频率位于最高频率阈值和最低频率阈值之间的情况下,则确定未发生异常,而目标接口的调用频率高于最高频率阈值或低于最低频率阈值的情况下,可以确定发生故障。

[0046] 方式二,通过目标接口的调用频率在频率数据集中的位置确定目标应用是否发生故障,频率数据集包括目标接口的调用频率。

[0047] 即,通过将当前的调用频率与历史大数据相比确定是否发生故障,如当前的调用频率位于历史大数据的范围内,则确定未发生故障,否则发生故障。

[0048] 例如,可以采用箱线图表示历史大数据。

[0049] 步骤131,获取历史大数据中目标接口的历史调用频率的最小值min和最大值max。

[0050] 步骤132,根据最小值和最大值确定第一四分位数Q1、中位数Q2以及第三四分位数Q3,第一四分位数为最大值max和最小值min的差值的四分之一与最小值之和,即Q1=min+(max-min)/4,中位数为最大值与最小值的和的二分之一,即Q2=(max+min)/2,第三四分位数为最大值和最小值的差值的四分之三与最小值之和,即Q3=min+(max-min)3/4。

[0051] 步骤133,获取第三四分位数与第一四分位数的差值为四分位全距IQR,即IQR=Q3-Q1。

[0052] 步骤134,获取第一四分位数与四分位全距的X倍之间的差值为频率下限(Q1-X*IQR),并获取第三四分位数与四分位全距的Y倍之间的和为频率上限(Q4+Y*IQR)。

[0053] 对于X和Y的取值,可以将多组数据代入上述公式,从而求取一个合适的X和Y的取值。

[0054] 在使用上述公式判断时,若目标接口的调用频率的在频率数据集的位置位于频率下限(Q1-X*IQR)与频率上限(Q4+Y*IQR)之间时,那么可以判断出应用侧是正常的,反之则发生故障。

[0055] 方式三,将相邻时间段的数据进行环比,一个时间周期可以分为多个时间段(即子周期),如将1天以6小时为单位可以划分为4个子周期(分别为0点至6点、6点到12点、12点到18点、18点到24点),通过比较第一子周期内的调用频率与第二子周期内的调用频率确定目标应用是否发生故障,第一子周期为当前获取的调用频率所在的时间子周期(如12点到18点),第二子周期为第一子周期的前一子周期(如6点到12点)。

[0056] 方式四,将当前时间段的数据与历史同期的数据进行同比,通过比较第一子周期内的调用频率与第三子周期内的调用频率确定目标应用是否发生故障,第一子周期在第一周期内的位置与第三子周期在第二周期内的位置相同,第二周期为第一周期的前一时间周期,例如,以天为时间周期,每天内子周期的划分方案与方式三中的方案相同,若第一周期为当天,则第二周期为当天的前一天,若第一子周期为当天的12点到18点,那么第三子周期即为前一天的12点到18点。

[0057] 在同比和环比的技术方案中,可以对比时间段内数据的走势、上下限等,若存在走势不同、上限过大或者下限过小,则确定发生故障,否则正常。

[0058] 步骤14,在基于调用数据确定服务系统是否发生服务故障之后或者过程中,可将当前获取到的目标接口的调用数据保存至数据库中,以便于以后作为历史大数据判断之后的数据是否异常,数据库可按照配置的过期时间将过期的调用数据删除。

[0059] 步骤15,在基于调用数据确定服务系统是否发生服务故障之后,向维护人员的终端推送报警,以便于对整个服务系统进行维护。

[0060] 相关技术中的日志监控告警是基于应用系统错误日志进行实时监控告警,未对接口调用频率进行主动监控。而本方案可在基于错误日志关键字进行实时监控告警的基础上新增调用频率监控,通过主动监控可以及时发现服务异常,而不是在反馈到业务数据报表上才察觉故障。

[0061] 基于错误日志的监控主要是监控服务本身,基于接口调用频率的监控不仅从调用频率角度监控了服务而且还监控了关联的上游服务调用方。扩大了监控范围,可作为错误日志监控的一种补充监控手段。特别是在微服务体系中,由于服务拆分比较多,扩大监控范围是非常有必要的,如图3所示,可将微服务的频率监控和错误日志监控进行整合。

[0062] 作为一种可选的实施例,下文结合具体的实施方案进一步详述本申请的技术方案:

[0063] 如图3所示,本方案可以整合频率监控和错误日志监控。如图4所示,错误日志监控侧重于系统内部错误日志的监控,侧重点在服务内部,如图5所示,调用频率监控侧重于服务暴露给外部接口的调用频率的监控,侧重点在外围。

[0064] 以上整合方案可以基于日志监控告警平台LogMonitor和日志web(英文全称World Wide Web,即全球广域网)在线监控工具普罗米修斯Prometheus实现,普罗米修斯的基本原理是通过HTTP(英文全称HyperText Transfer Protocol,即超文本传输协议)协议周期性

抓取被监控组件的状态,任意组件只要提供对应的HTTP接口就可以接入监控,输出被监控组件信息的HTTP接口被叫做指标采集接口exporter(exporter是普罗米修斯监控中重要的组成部分,负责数据指标的采集)。目前互联网公司常用的组件大部分都有指标采集接口可以直接使用。

[0065] 如图5所示,可通过普罗米修斯客户端收集指定时间段各接口的调用数,从而计算各个维度的调用频率。每个服务接口可暴露一个HTTP服务的接口给普罗米修斯,以便于进行频率数据的定时抓取。普罗米修斯支持通过配置文件、文本文件等方式预先配置抓取的目标,如需要抓取的数据所在的接口、该数据所在的字段等;配置完成后,普罗米修斯定时去接口上抓取频率数,可在本地存储抓取的所有频率数据,并通过一定规则进行清理和整理数据,并把得到的结果存储到新的时间序列中。

[0066] 在上述方案中,普罗米修斯的频率监控流程如图6所示:

[0067] 步骤21,普罗米修斯的服务组件(prometheus server)通过HTTP协议从服务接口中周期性的拉取应用指标数据,如指定周期内每分钟的各个接口的请求数。

[0068] 步骤22,按照业务特性和以往数据指标呈现指定频率规则,遍历已开启的频率规则,按规则配置时间定期拉取指标,然后存储在数据库redis,redis数据的过期时间由指标的对比周期来决定。

[0069] 如同比月数据,数据过期时间设置为大于两个月即可,同比年数据则过期时间设置则要大于两年,设置过期时间是为了避免指标数据量太多而把缓存数据库塞满,当数据不再使用后,即到了过期时间就自动清理。

[0070] 对比周期是由日志监控告警平台配置的监控规则来定的,如接口A的频率统计周期是天,那只需要缓存最近一个月的指标数据即可,通过箱线法将当天统计的指标数据和最近一个月的做比对,判定是否异常。如果是一个月,那就需要缓存大概一年的对比指标数据。

[0071] 步骤23,日志监控告警平台触发指标对比时,先从普罗米修斯中获取指标。

[0072] 此处指标为按照为系统接口调用预设的指标采集规则进行采集得到的指标,是采集统计的系统指标,后续用此指标数据和阈值做对比,以判断是否异常。

[0073] 步骤24,将频率指标对比人工预设阈值,将指标存入redis,与预设阈值对比,返回对比结果。

[0074] 如图7所示,是一个生产环境活动的主页接口在一段时间内的调用量,横轴表示时间、纵轴表示频率,假如其调用频率每分钟在200次以上,对于此接口可以设定调用频率阈值100/min,低于100可认为是异常的,对于频率波动大的指标也可以引入箱线图方法。

[0075] 步骤25,对于采用同比和环比的方案,可将指标存入redis,然后拉取同比或者环比对象,进行对比,对比结果与预设值进行比较,返回最终对比结果。

[0076] 具体是采用同比还是环比,可根据业务接口特性确定,如同比可以是本月第一周和上个月第一周的对比,也可能是今年这个月和去年这个月对比;环比可以是本周和上周的对比,也可以是本月和上月的对比。

[0077] 步骤26,判断对比结果是否达到告警级别,如果达到则触发响应级别的告警。以采用箱线图的告警为例:

[0078] 如图8所示的箱线图,其计算步骤如下:

[0079] 找出第一个四分位置Q1、中位数Q2、第三个四分位置Q3；

[0080] 计算四分位全距:IQR＝Q3-Q1；

[0081] 计算上下限,例如为Q1-1.5IQR、Q3+1.5IQR；

[0082] 由于各个系统的各个被调用的接口有自己的业务特性,可跟进实际情况设定系数,Q1-X*IQR≤T≤Q3+Y*IQR,X、Y的取值可根据实际业务情况配置的调控参数持续调整优化,提供告警的准确度。

[0083] 对于以天为统计调用周期的接口,取60天的数据,分为5组12天的调用数据,以此确定箱线图的比例系数,如果近12个天中有异常点数据被剔除,则继续往前取一位。

[0084] 取到的一组数并按顺序排列:X1、X2、X3、X4、X5、X6、X7、X8、X9、X10、X11、X12。

[0085] 查找第一个四分位:Q1＝(12+1)/4＝3,则Q1＝X3；

[0086] Q3＝3*(12+1)/4＝9,则Q3＝X9；

[0087] IQR＝X9-X3；

[0088] X3-X(X9-X3)≤T≤X9+Y(X9-X3)；

[0089] 通过五组数据,根据临界点求出箱线图的比例系数临界点X,和临界点Y。

[0090] 随着数据的生成重新计算动态调整X和Y,对于当日调用统计结果T,判断T是否满足条件X3-m*X(X9-X3)≤T≤X9+n*Y(X9-X3),其中m,n是可调控系数,具体设定的大小根据接口调用频率波动来设定。

[0091] 本方案通过接口频率的监控可以监控本系统接口被调用的健康情况,扩展了监控范围,使服务本身及上游服务也获得监控保护。对系统频率数据进行实时分析,系统自动调节各自对比指标场景下的监控阈值参数,同时系统可通过自身接口数据分析自动调节频率拉取周期,以获得更准确的监控结果。

[0092] 将上报到普罗米修斯的接口的流量,通过时间维度统计调用频率。然后根据业务实际特性进行同比,环比等方式进行确定设置接口调用频率告警阈值。当调用低于阈值时进行电话或者其他方式的告警。此方案可对外部服务故障导致调用(回调)当前服务接口频率降低,或者中断的情况进行有效监控。

[0093] 需要说明的是,对于前述的各方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本申请并不受所描述的动作顺序的限制,因为依据本申请,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本申请所必须的。

[0094] 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根据上述实施例的方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,或者网络设备等)执行本申请各个实施例所述的方法。

[0095] 根据本申请实施例的另一个方面,还提供了一种用于实施上述服务的监控方法的服务的监控装置。图9是根据本申请实施例的一种可选的服务的监控装置的示意图,如图9所示,该装置可以包括：

[0096] 接收单元901,用于接收监控指令,其中,所述监控指令用于指示监控服务系统是否发生服务异常;

[0097] 获取单元903,用于响应于所述监控指令,获取服务接口的调用数据,其中,所述调用数据为调用端的应用通过所述服务接口调用服务端的服务时产生的数据,所述服务系统包括所述调用端、所述服务接口以及所述服务端;

[0098] 监控单元905,用于基于所述调用数据确定所述服务系统是否发生服务故障。

[0099] 需要说明的是,该实施例中的接收单元901可以用于执行本申请实施例中的步骤S202,该实施例中的获取单元903可以用于执行本申请实施例中的步骤S204,该实施例中的监控单元905可以用于执行本申请实施例中的步骤S206。

[0100] 此处需要说明的是,上述模块与对应的步骤所实现的示例和应用场景相同,但不限于上述实施例所公开的内容。需要说明的是,上述模块作为装置的一部分可以运行在如图1所示的硬件环境中,可以通过软件实现,也可以通过硬件实现。

[0101] 利用日志分析故障,需要进行问题累计,累计完成后在需要检测故障时,收集相对于时间而言具有一定信息延迟的日志,然后通过对日志的分析确定是否发生故障,其中的问题累计和日志分析需要消耗时间,且分析的日志所记载的信息具有一定的延时性,从而造成故障检测的时效性较差,而在本申请的技术方案中,通过上述方案,在需要监控服务系统是否发生服务异常时,可获取服务接口的调用数据,基于调用数据确定服务系统是否发生服务故障,由于获取的调用数据是实时数据,进而可以在第一时间确定是否发生故障,可以解决相关技术中的故障检测时效性较差的技术问题,进而达到提高检测时效性的技术效果。

[0102] 可选地,获取单元还用于获取多个所述服务接口的调用数据;监控单元还可用于:根据多个所述服务接口中每个所述服务接口的调用数据确定所述服务接口、位于所述服务接口的调用端的应用以及位于所述服务接口的服务端的服务中的至少之一是否发生故障。

[0103] 可选地,监控单元还可用于:在目标接口收到的请求数据发生异常的情况下,确定目标应用发生故障,其中,所述目标接口为当前处理的服务接口,所述目标应用为位于所述目标接口的调用端的应用,所述调用数据包括所述请求数据;在所述目标接口收到的请求数据正常且与请求数据对应的响应数据发生异常的情况下,确定目标服务发生异常,其中,所述目标服务为位于所述目标接口的服务端的服务,所述调用数据包括所述响应数据;在所述目标接口对调用数据的传输频率低于第一阈值且与提供所述目标服务的其他服务接口对调用数据的传输频率高于第二阈值的情况下,确定所述目标接口发生异常,其中,所述第二阈值高于所述第一阈值。

[0104] 可选地,监控单元还可用于:将产生当前待处理的调用数据的所述服务接口作为目标接口;通过分析所述目标接口的调用数据确定所述目标接口、目标应用以及目标服务中的至少之一是否发生故障,其中,所述目标应用为位于所述目标接口的调用端的应用,所述目标服务为位于所述目标接口的服务端的服务。

[0105] 可选地,所述调用数据包括调用频率,监控单元还可用于:通过比较所述目标接口的调用频率和频率阈值确定所述目标应用是否发生故障;通过所述目标接口的调用频率在频率数据集中的位置确定所述目标应用是否发生故障,其中,所述频率数据集包括所述目标接口的调用频率;通过比较第一子周期内的调用频率与第二子周期内的调用频率确定所

述目标应用是否发生故障,其中,所述第一子周期为当前获取的调用频率所在的时间子周期,所述第二子周期为所述第一子周期的前一子周期;通过比较所述第一子周期内的调用频率与第三子周期内的调用频率确定所述目标应用是否发生故障,其中,所述第一子周期在第一周期内的位置与所述第三子周期在第二周期内的位置相同,所述第二周期为所述第一周期的前一时间周期。

[0106] 可选地,监控单元还可用于:在通过比较所述目标接口的调用频率和频率阈值确定所述目标应用是否发生故障之前,统计所述服务系统正常运行时所述目标接口的最低调用频率和最高调用频率;将所述最低调用频率与缩小系数之间的乘积作为所述最低频率阈值,将所述最高调用频率与放大系数之间的乘积作为所述最高频率阈值。

[0107] 可选地,监控单元还可用于:在通过所述目标接口的调用频率在频率数据集中的位置确定所述目标应用是否发生故障之前,获取所述目标接口的历史调用频率的最小值和最大值;根据所述最小值和所述最大值确定第一四分位数Q1、中位数Q2以及第三四分位数Q3,其中,所述第一四分位数为所述最大值和所述最小值的差值的四分之一与所述最小值之和,所述中位数为所述最大值与所述最小值的和的二分之一,所述第三四分位数为所述最大值和所述最小值的差值的四分之三与所述最小值之和;获取所述第三四分位数与所述第一四分位数的差值为四分位全距IQR;获取所述第一四分位数与所述四分位全距的X倍之间的差值为频率下限(Q1-X*IQR),并获取第三四分位数与所述四分位全距的Y倍之间的和为频率上限(Q4+Y*IQR)。

[0108] 可选地,监控单元还可用于:在基于所述调用数据确定所述服务系统是否发生服务故障之后,将当前获取到的目标接口的调用数据保存至数据库中,其中,所述数据库按照配置的过期时间将过期的调用数据删除。

[0109] 本方案通过接口频率的监控可以监控本系统接口被调用的健康情况,扩展了监控范围,使服务本身及上游服务也获得监控保护。对系统频率数据进行实时分析,系统自动调节各自对比指标场景下的监控阈值参数,同时系统可通过自身接口数据分析自动调节频率拉取周期,以获得更准确的监控结果。

[0110] 将上报到普罗米修斯的接口的流量,通过时间维度统计调用频率。然后根据业务实际特性进行同比,环比等方式进行确定设置接口调用频率告警阈值。当调用低于阈值时进行电话或者其他方式的告警。此方案可对外部服务故障导致调用(回调)当前服务接口频率降低,或者中断的情况进行有效监控。

[0111] 此处需要说明的是,上述模块与对应的步骤所实现的示例和应用场景相同,但不限于上述实施例所公开的内容。需要说明的是,上述模块作为装置的一部分可以运行在如图1所示的硬件环境中,可以通过软件实现,也可以通过硬件实现,其中,硬件环境包括网络环境。

[0112] 根据本申请实施例的另一个方面,还提供了一种用于实施上述服务的监控方法的服务器或终端。

[0113] 图10是根据本申请实施例的一种终端的结构框图,如图10所示,该终端可以包括:一个或多个(图中仅示出一个)处理器1001、存储器1003、以及传输装置1005,如图10所示,该终端还可以包括输入输出设备1007。

[0114] 其中,存储器1003可用于存储软件程序以及模块,如本申请实施例中的服务的监

控方法和装置对应的程序指令/模块,处理器1001通过运行存储在存储器1003内的软件程序以及模块,从而执行各种功能应用以及数据处理,即实现上述的服务的监控方法。存储器1003可包括高速随机存储器,还可以包括非易失性存储器,如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中,存储器1003可进一步包括相对于处理器1001远程设置的存储器,这些远程存储器可以通过网络连接至终端。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0115] 上述的传输装置1005用于经由一个网络接收或者发送数据,还可以用于处理器与存储器之间的数据传输。上述的网络具体实例可包括有线网络及无线网络。在一个实例中,传输装置1005包括一个网络适配器(Network Interface Controller,NIC),其可通过网线与其他网络设备与路由器相连从而可与互联网或局域网进行通讯。在一个实例中,传输装置1005为射频(Radio Frequency,RF)模块,其用于通过无线方式与互联网进行通讯。

[0116] 其中,具体地,存储器1003用于存储应用程序。

[0117] 处理器1001可以通过传输装置1005调用存储器1003存储的应用程序,以执行下述步骤:

[0118] 接收监控指令,其中,所述监控指令用于指示监控服务系统是否发生服务异常;

[0119] 响应于所述监控指令,获取服务接口的调用数据,其中,所述调用数据为调用端的应用通过所述服务接口调用服务端的服务时产生的数据,所述服务系统包括所述调用端、所述服务接口以及所述服务端;

[0120] 基于所述调用数据确定所述服务系统是否发生服务故障。

[0121] 处理器1001还用于执行下述步骤:

[0122] 通过比较所述目标接口的调用频率和频率阈值确定所述目标应用是否发生故障;

[0123] 通过所述目标接口的调用频率在频率数据集中的位置确定所述目标应用是否发生故障,其中,所述频率数据集包括所述目标接口的调用频率;

[0124] 通过比较第一子周期内的调用频率与第二子周期内的调用频率确定所述目标应用是否发生故障,其中,所述第一子周期为当前获取的调用频率所在的时间子周期,所述第二子周期为所述第一子周期的前一子周期;

[0125] 通过比较所述第一子周期内的调用频率与第三子周期内的调用频率确定所述目标应用是否发生故障,其中,所述第一子周期在第一周期内的位置与所述第三子周期在第二周期内的位置相同,所述第二周期为所述第一周期的前一时间周期。

[0126] 相关技术中,利用日志分析故障,需要进行问题累计,累计完成后在需要检测故障时,收集相对于时间而言具有一定信息延迟的日志,然后通过对日志的分析确定是否发生故障,其中的问题累计和日志分析需要消耗时间,且分析的日志所记载的信息具有一定的延时性,从而造成故障检测的时效性较差,而在本申请的技术方案中,通过上述步骤,在需要监控服务系统是否发生服务异常时,可获取服务接口的调用数据,基于所述调用数据确定所述服务系统是否发生服务故障,由于获取的调用数据是实时数据,进而可以在第一时间确定是否发生故障,可以解决相关技术中的故障检测时效性较差的技术问题,进而达到提高检测时效性的技术效果。

[0127] 可选地,本实施例中的具体示例可以参考上述实施例中所描述的示例,本实施例在此不再赘述。

[0128] 本领域普通技术人员可以理解,图10所示的结构仅为示意,终端可以是智能手机(如Android手机、iOS手机等)、平板电脑、掌上电脑以及移动互联网设备(Mobile Internet Devices,MID)、PAD等终端设备。图10其并不对上述电子装置的结构造成限定。例如,终端还可包括比图10中所示更多或者更少的组件(如网络接口、显示装置等),或者具有与图10所示不同的配置。

[0129] 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通过程序来指令终端设备相关的硬件来完成,该程序可以存储于一计算机可读存储介质中,存储介质可以包括:闪存盘、只读存储器(Read-Only Memory,ROM)、随机存取器(Random Access Memory,RAM)、磁盘或光盘等。

[0130] 本申请的实施例还提供了一种存储介质。可选地,在本实施例中,上述存储介质可以用于执行服务的监控方法的程序代码。

[0131] 可选地,在本实施例中,上述存储介质可以位于上述实施例所示的网络中的多个网络设备中的至少一个网络设备上。

[0132] 可选地,在本实施例中,存储介质被设置为存储用于执行以下步骤的程序代码:

[0133] 接收监控指令,其中,所述监控指令用于指示监控服务系统是否发生服务异常;

[0134] 响应于所述监控指令,获取服务接口的调用数据,其中,所述调用数据为调用端的应用通过所述服务接口调用服务端的服务时产生的数据,所述服务系统包括所述调用端、所述服务接口以及所述服务端;

[0135] 基于所述调用数据确定所述服务系统是否发生服务故障。

[0136] 可选地,存储介质还被设置为存储用于执行以下步骤的程序代码:

[0137] 通过比较所述目标接口的调用频率和频率阈值确定所述目标应用是否发生故障;

[0138] 通过所述目标接口的调用频率在频率数据集中的位置确定所述目标应用是否发生故障,其中,所述频率数据集包括所述目标接口的调用频率;

[0139] 通过比较第一子周期内的调用频率与第二子周期内的调用频率确定所述目标应用是否发生故障,其中,所述第一子周期为当前获取的调用频率所在的时间子周期,所述第二子周期为所述第一子周期的前一子周期;

[0140] 通过比较所述第一子周期内的调用频率与第三子周期内的调用频率确定所述目标应用是否发生故障,其中,所述第一子周期在第一周期内的位置与所述第三子周期在第二周期内的位置相同,所述第二周期为所述第一周期的前一时间周期。

[0141] 可选地,本实施例中的具体示例可以参考上述实施例中所描述的示例,本实施例在此不再赘述。

[0142] 可选地,在本实施例中,上述存储介质可以包括但不限于:U盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

[0143] 根据本申请的一个方面,提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设备执行上述方案的各种可选实现方式中提供的方法。

[0144] 上述本申请实施例序号仅仅为了描述,不代表实施例的优劣。

[0145] 上述实施例中的集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时,可以存储在上述计算机可读取的存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在存储介质中,包括若干指令用以使得一台或多台计算机设备(可为个人计算机、服务器或者网络设备等)执行本申请各个实施例所述方法的全部或部分步骤。

[0146] 在本申请的上述实施例中,对各个实施例的描述都各有侧重,某个实施例中没有详述的部分,可以参见其他实施例的相关描述。

[0147] 在本申请所提供的几个实施例中,应该理解到,所揭露的客户端,可通过其它的方式实现。其中,以上所描述的装置实施例仅仅是示意性的,例如所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,单元或模块的间接耦合或通信连接,可以是电性或其它的形式。

[0148] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0149] 另外,在本申请各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0150] 以上所述仅是本申请的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本申请原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本申请的保护范围。

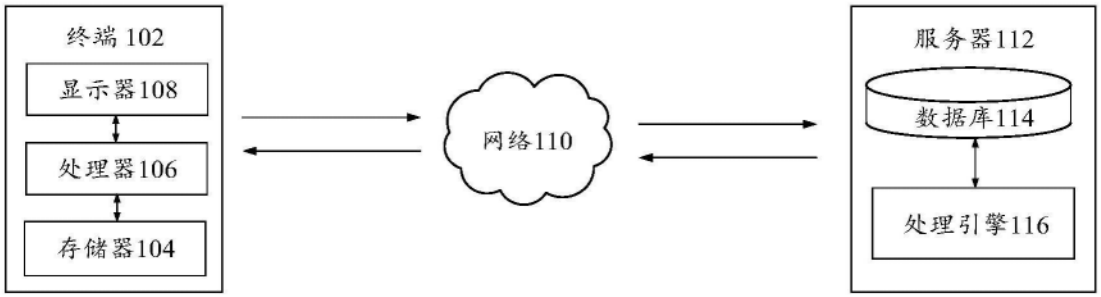

终端 102
显示器108
处理器106
存储器104
网络110
服务器112
数据库114
处理引擎116

图1

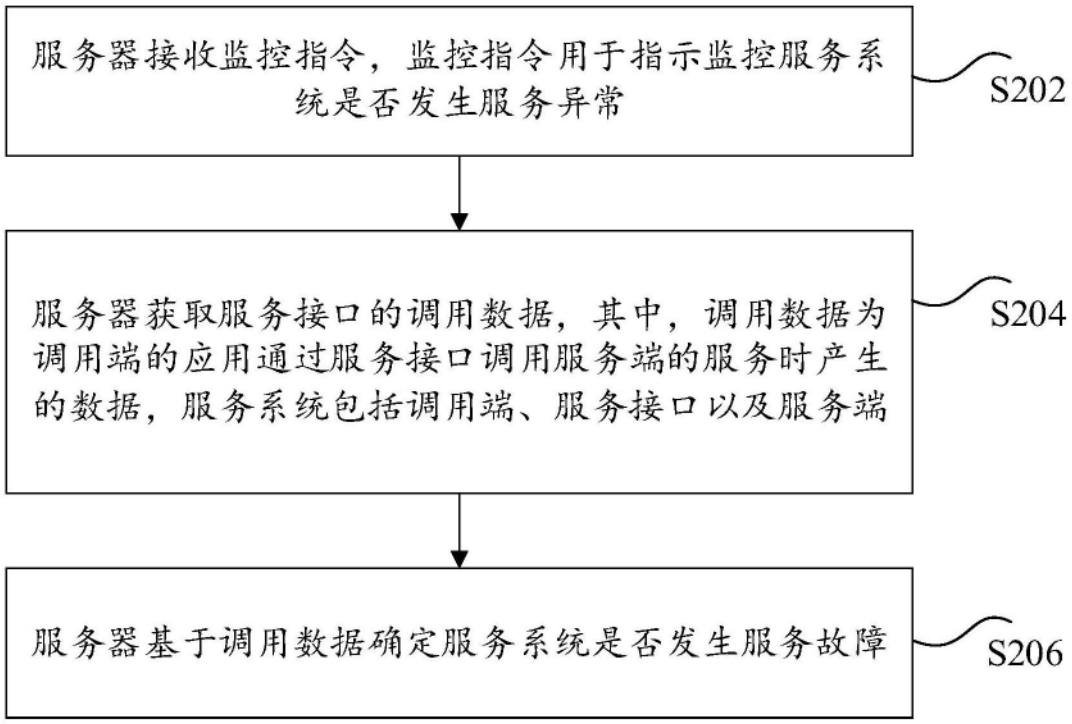

服务器接收监控指令，监控指令用于指示监控服务系统是否发生服务异常
S202
服务器获取服务接口的调用数据，其中，调用数据为调用端的应用通过服务接口调用服务端的服务时产生的数据，服务系统包括调用端、服务接口以及服务端
S204
服务器基于调用数据确定服务系统是否发生服务故障
S206

图2

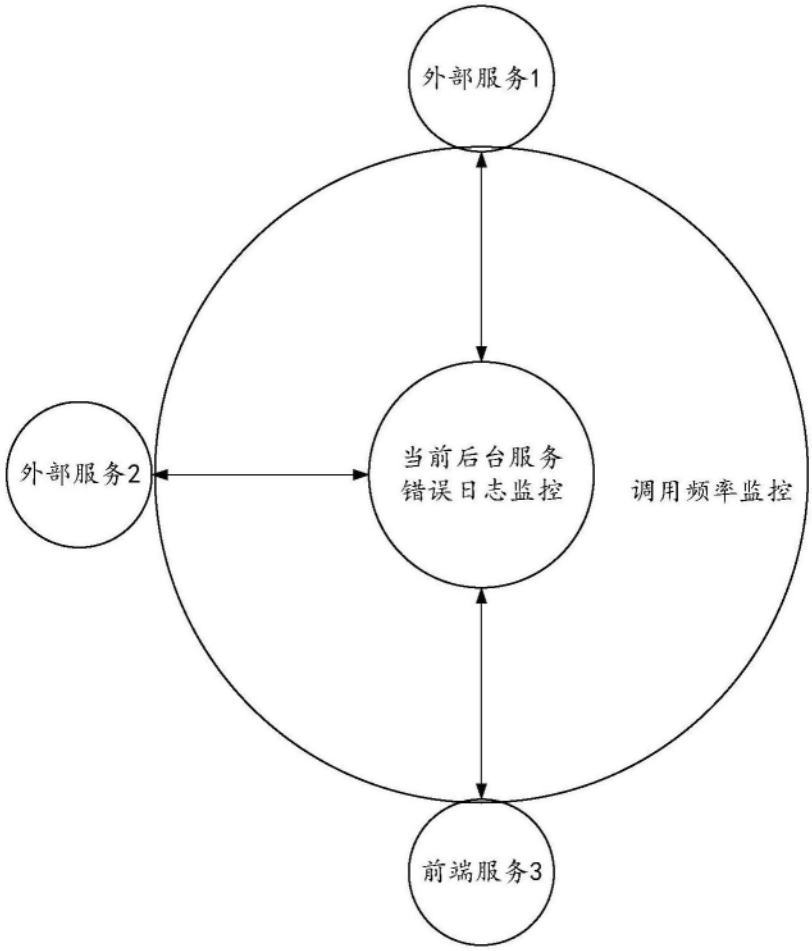


图3

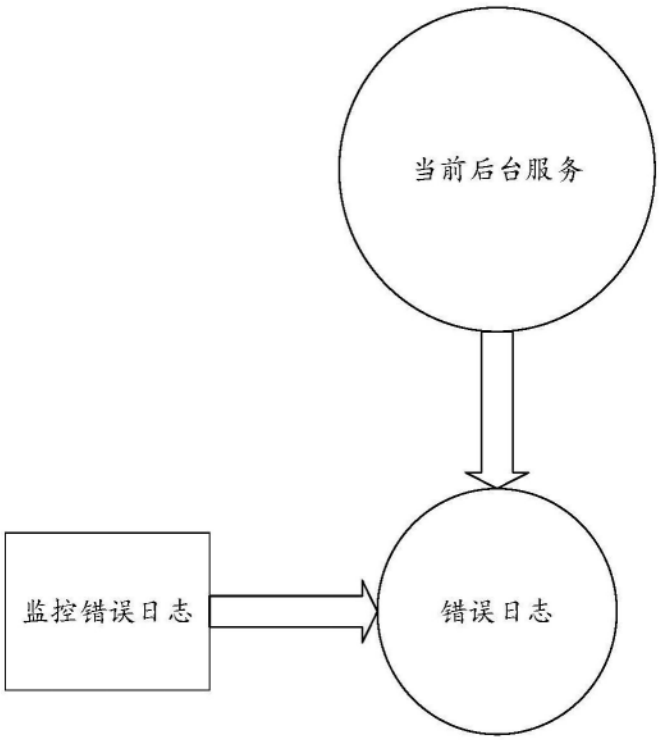


图4

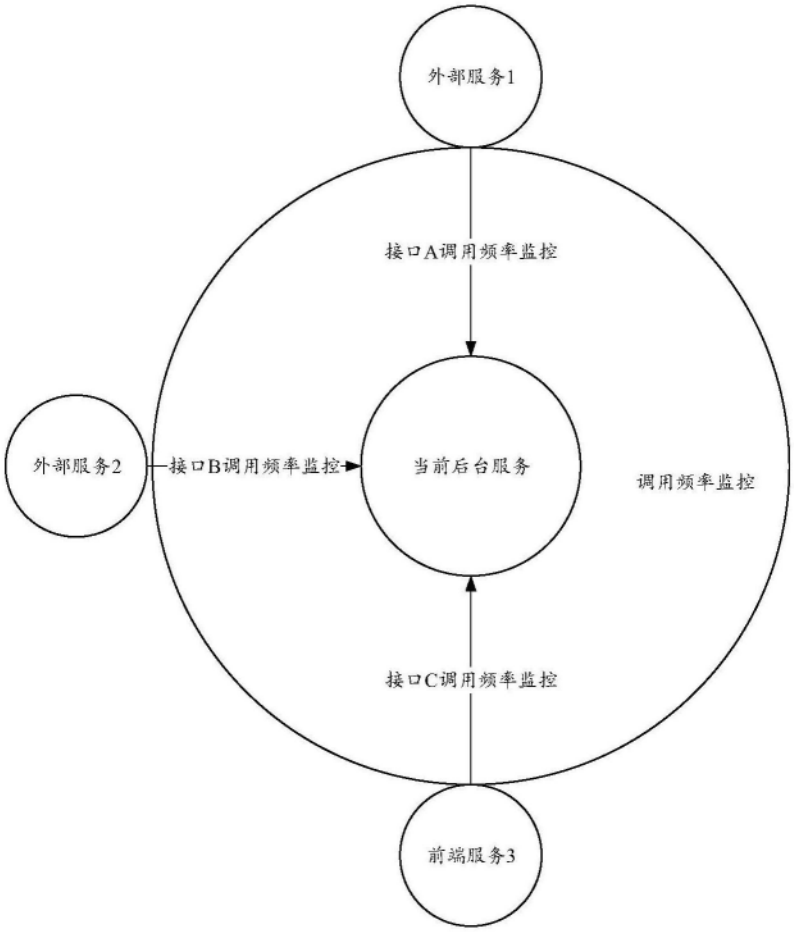

外部服务1
接口A调用频率监控
外部服务2
接口B调用频率监控
当前后台服务
调用频率监控
接口C调用频率监控
前端服务3

图5

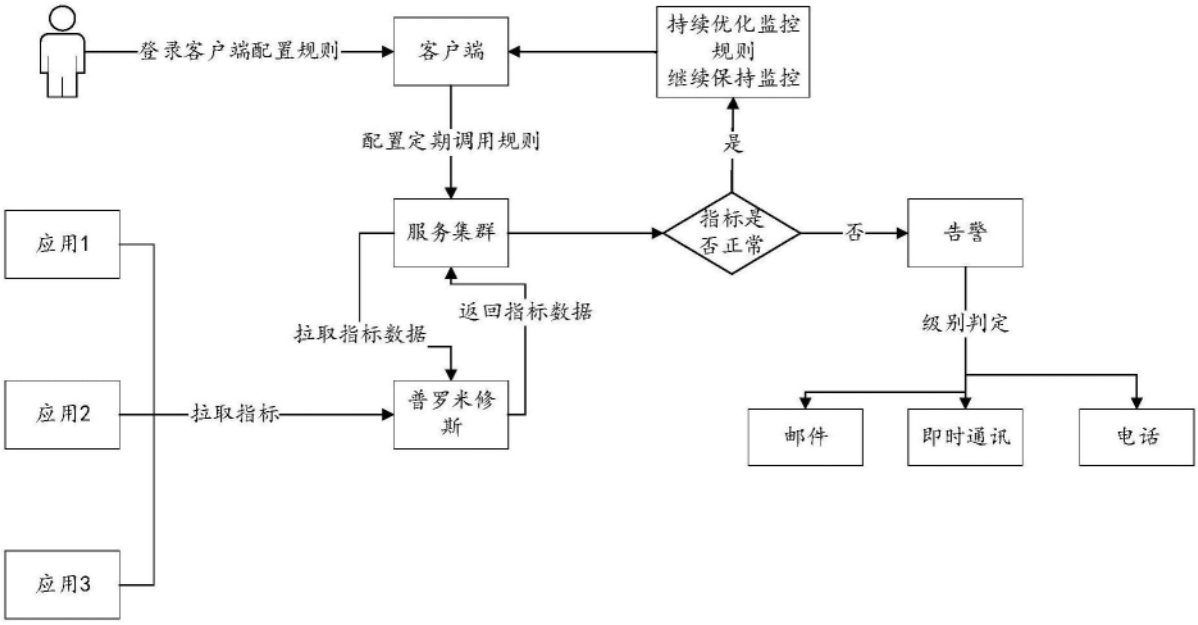

登录客户端配置规则
客户端
持续优化监控规则
继续保持监控
配置定期调用规则
是
应用1
服务集群
指标是否正常
否
告警
拉取指标数据
返回指标数据
级别判定
应用2
拉取指标
普罗米修斯
邮件
即时通讯
电话
应用3

图6

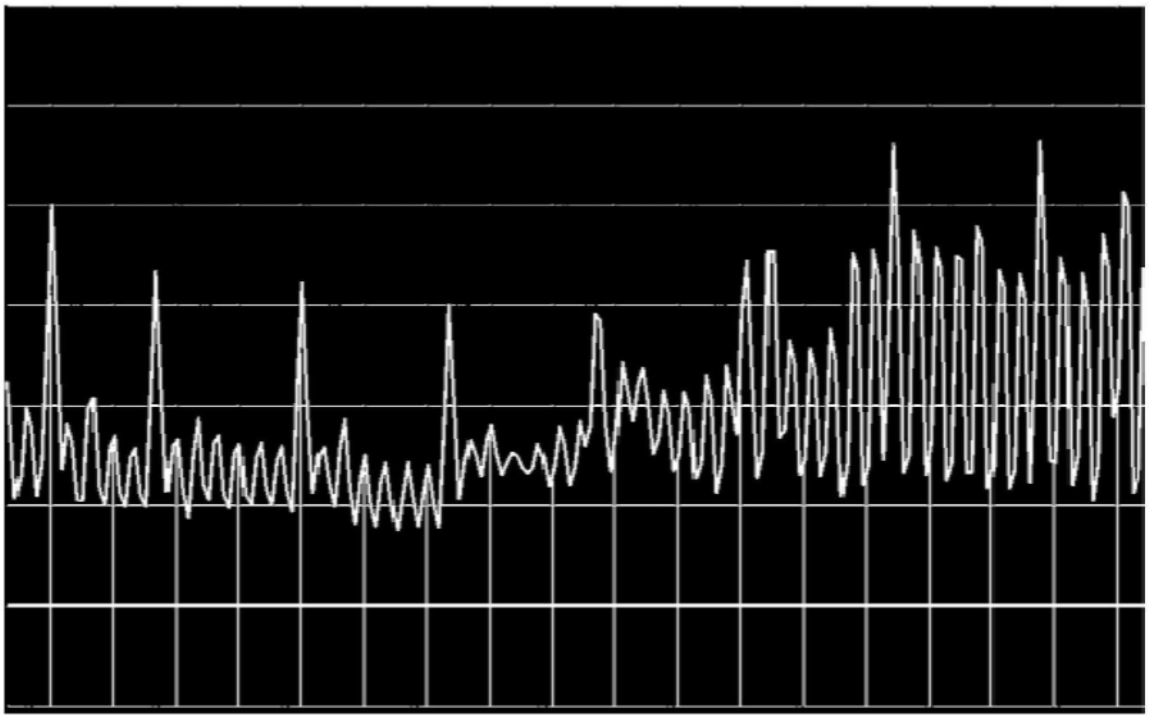


图7

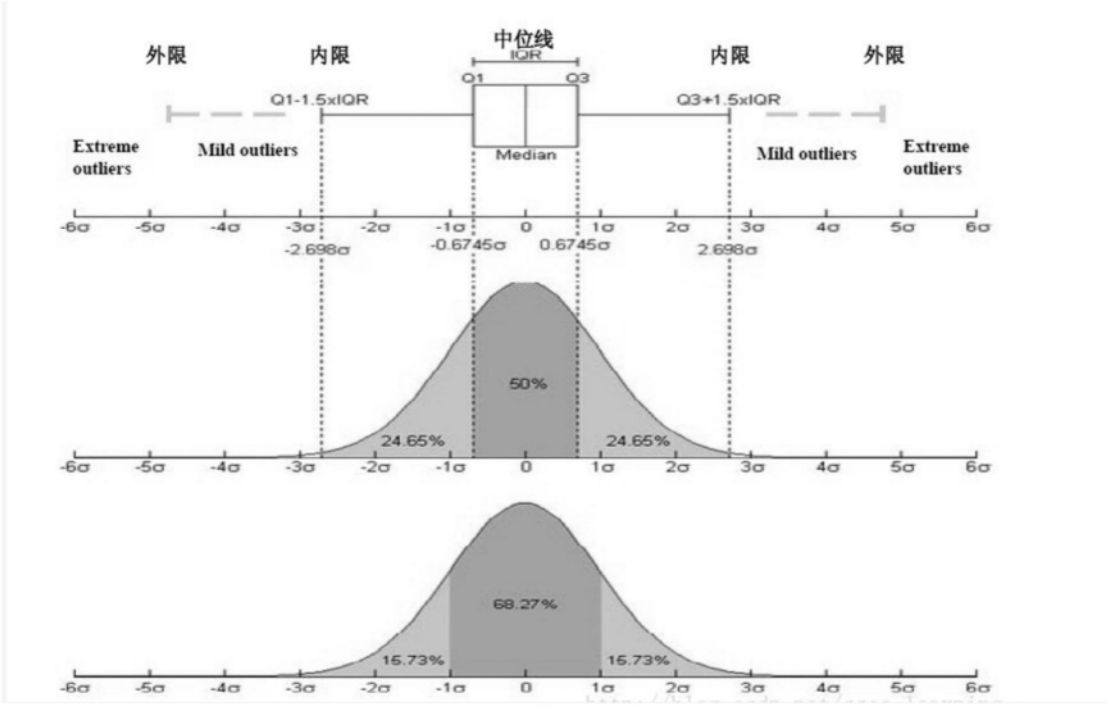

外限
内限
中位线
IQR
内限
外限
Q1
Q3
Q1-1.5xIQR
Q3+1.5xIQR
Extreme outliers
Mild outliers
Median
Mild outliers
Extreme outliers
-6σ -5σ -4σ -3σ -2σ -1σ 0 1σ 2σ 3σ 4σ 5σ 6σ
-2.698σ
-0.6745σ
0.6745σ
2.698σ
50%
24.65%
24.65%
68.27%
15.73%
15.73%

图8

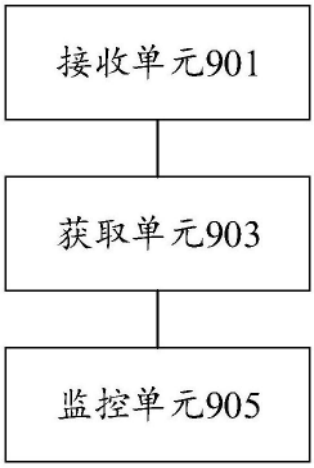


图9

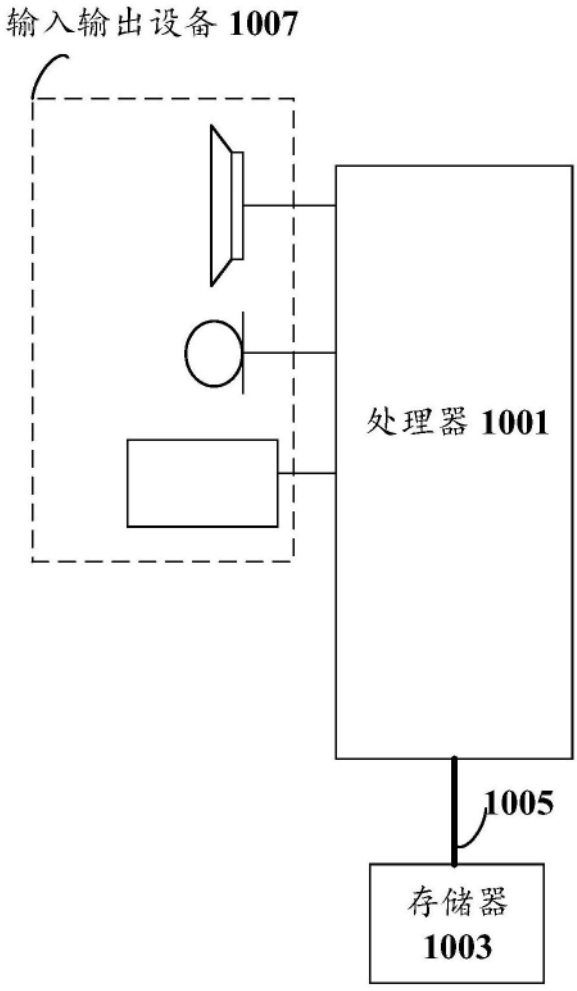


图10