(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 110138732 B
(45)授权公告日 2022.03.29

(21)申请号 201910268069.5

(22)申请日 2019.04.03

(65)同一申请的已公布的文献号
申请公布号 CN 110138732 A

(43)申请公布日 2019.08.16

(73)专利权人 平安科技(深圳)有限公司
地址 518000 广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼

(72)发明人 朱坤

(74)专利代理机构 深圳市力道知识产权代理事务所(普通合伙) 44507
代理人 何姣

(51)Int.Cl.
H04L 9/40(2022.01)
H04L 67/1042(2022.01)
H04L 67/1008(2022.01)
H04L 67/1025(2022.01)

(56)对比文件
CN 103581039 A,2014.02.12
CN 107613030 A,2018.01.19
CN 102394931 A,2012.03.28
CN 103825785 A,2014.05.28
CN 109347982 A,2019.02.15
CN 105100237 A,2015.11.25
CN 105808347 A,2016.07.27
CN 106959894 A,2017.07.18
CN 109218100 A,2019.01.15
CN 108667878 A,2018.10.16

审查员 谭雪

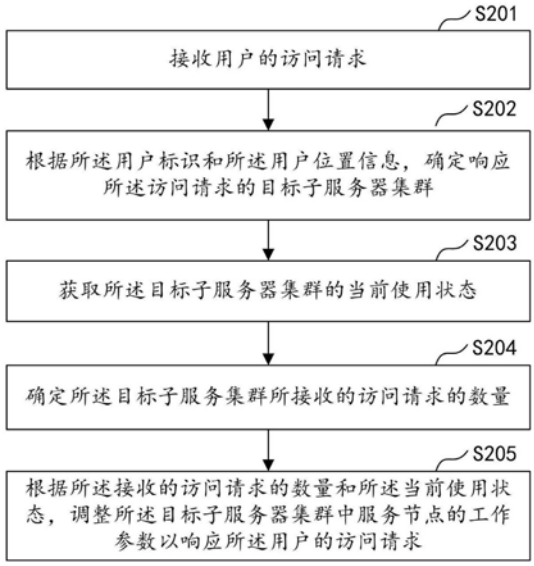
权利要求书4页 说明书14页 附图7页

(54)发明名称

访问请求的响应方法、装置、设备及存储介质

(57)摘要

本申请涉及云计算领域，具体使用了所述接收的访问请求的数量和目标子服务集群的当前使用状态实现负载调配，并公开了一种访问请求的响应方法、装置、设备及存储介质，所述方法包括：接收用户的访问请求，所述访问请求包含用户标识和用户位置信息；根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群；获取所述目标子服务器集群的当前使用状态；确定所述目标子服务集群所接收的访问请求的数量；根据所述接收的访问请求的数量和所述当前使用状态，调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求。进而动态调整服务器集群的服务节点的工作状态，提高服务器集群的响应速度，由此提高用户的访问体验度。



CN 110138732 B

1.一种访问请求的响应方法,其特征在于,包括:

接收用户的访问请求,所述访问请求包含用户标识和用户位置信息;

根据所述用户标识和所述用户位置信息,确定响应所述访问请求的目标子服务器集群;

获取所述目标子服务器集群的当前使用状态;

确定所述目标子服务集群所接收的访问请求的数量;

根据所述接收的访问请求的数量和所述当前使用状态,调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求;

其中,根据所述接收的访问请求的数量和所述当前使用状态,调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求,包括:

根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态,计算所述目标子服务器集群中当前启动的服务节点的资源使用率;

根据所述目标子服务集群的总用户请求量、总请求处理能力和总请求处理强度,计算所述目标子服务集群的总资源使用率;

若所述资源使用率大于第一阀值且所述总资源使用率大于第三阀值时,则在目标子服务器集群中启动新的服务节点;若所述资源使用率大于所述第一阀值且所述总资源使用率小于第四阀值时,则拒绝接收新的访问请求;若所述资源使用率小于第二阀值时,则在目标子服务器集群中关闭一个当前启动的服务节点,所述第二阀值小于所述第一阀值,所述第四阀值小于所述第三阀值。

2.根据权利要求1所述的访问请求的响应方法,其特征在于,所述根据所述用户标识和所述用户位置信息,确定响应所述访问请求的目标子服务器集群,包括:

根据所述用户标识获取用户的注册地址;

根据所述注册地址和所述用户位置信息确定响应所述访问请求的目标子服务集群。

3.根据权利要求1所述的访问请求的响应方法,其特征在于,所述根据所述接收的访问请求的数量和所述当前使用状态,调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求,包括:

基于使用率计算公式,根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态,计算所述目标子服务器集群中当前启动的服务节点的资源使用率;

若所述资源使用率大于第一阀值时,则在所述目标子服务器集群中启动新的服务节点;若所述资源使用率小于第二阀值时,则在所述目标子服务器集群中关闭一个当前启动的服务节点;其中,所述第二阀值小于所述第一阀值。

4.根据权利要求3所述的访问请求的响应方法,其特征在于,所述使用率计算公式为:

$$T_i=\begin{cases}1, & if(r_i=R_i \ or \ q_i=Q_i \ or \ s_i=S_i)\\ w_1\dfrac{r_i}{R_i}+w_2\dfrac{q_i}{Q_i}+w_3\dfrac{s_i}{S_i}, & Others\end{cases}$$

其中,T_i表示目标子服务集群中当前启动的第i个服务节点的资源使用率,r_i、R_i分别表示服务节点的用户请求量和最大请求量,q_i、Q_i分别表示服务节点的请求处理能力和最大处

理能力，s_i、S_i分别表示服务节点的请求处理强度和最大处理强度，$S_i=\frac{r_i}{C_i \cdot q_i}$，$C_i$为目标子服务集群中当前启动的第i个服务节点的访问请求的并行处理能力，w_1、w_2、w_3分别表示用户请求量、请求处理能力和请求处理强度的权重，且$w_1+w_2+w_3=1$。

5.根据权利要求3或4所述的访问请求的响应方法，其特征在于，所述基于使用率计算公式，根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，计算所述目标子服务器集群中当前启动的服务节点的资源使用率，包括：

基于总使用率计算公式，根据所述目标子服务集群的总用户请求量、总请求处理能力和总请求处理强度，计算所述目标子服务集群的总资源使用率；

若所述资源使用率大于所述第一阀值且所述总资源使用率大于第三阀值时，则在目标子服务器集群中启动新的服务节点；若所述资源使用率大于所述第一阀值且所述总资源使用率小于第四阀值时，则拒绝接收新的访问请求；若所述资源使用率小于所述第二阀值时，则在目标子服务器集群中关闭一个当前启动的服务节点；其中，所述第四阀值小于所述第三阀值；

所述总使用率计算公式为：

$$T_N=\begin{cases}1, & if(r_N=R_N \ or \ q_N=Q_N \ or \ s_N=S_N)\\ w_4\frac{r_N}{R_N}+w_5\frac{q_N}{Q_N}+w_6\frac{s_N}{S_N}, & Others\end{cases}$$

其中，T_N表示目标子服务集群的总资源使用率，r_N、R_N分别表示目标子服务集群的总用户请求量和最大总请求量，q_N、Q_N分别表示目标子服务集群的总请求处理能力和最大总处理能力，s_N、S_N分别表示目标子服务集群的总请求处理强度和最大总请求处理强度，w_4，w_5，w_6分别表示总用户请求量、总请求处理能力和总请求处理强度的权重，且$w_4+w_5+w_6=1$。

6.根据权利要求5所述的访问请求的响应方法，其特征在于，所述基于总使用率计算公式，根据所述目标子服务集群的总用户请求量、总请求处理能力和总请求处理强度，计算所述目标子服务集群的总资源使用率之前，还包括：

基于总请求量计算公式，根据所述目标子服务集群的当前启动的各服务节点的用户请求量，计算所述目标子服务集群的总用户请求量；

基于总处理能力计算公式，根据所述目标子服务集群的当前启动的各服务节点的请求处理能力，计算所述目标子服务集群的总请求处理能力；

基于总处理强度计算公式，根据所述总用户请求量和总请求处理能力，计算所述目标子服务集群的总请求处理强度；

所述总请求量计算公式为：

$$r_N=\sum_{i=1}^{n}r_i$$

其中，r_N表示目标子服务集群的总用户请求量，r_i表示目标子服务集群的当前启动的第i个服务节点的用户请求量，n为目标子服务集群中当前启动的服务节点的数量；

所述总处理能力计算公式为：

$$q_N = n / \sum_{i=1}^{n} \frac{1}{q_i}$$

其中，q_N表示目标子服务集群的总请求处理能力，q_i表示目标子服务集群的当前启动的第i个服务节点的请求处理能力；

所述总处理强度计算公式为：

$$s_N = \frac{r_N}{(q_N \cdot \sum_{i=1}^{n} C_i)}$$

其中，s_N表示目标子服务集群的总请求处理强度，C_i为目标子服务集群中当前启动的第i个服务节点的访问请求的并行处理能力。

7.根据权利要求1所述的访问请求的响应方法，其特征在于，所述根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求之后，还包括：

判断在预设时间内所述目标子服务集群是否接收到新的访问请求；

若在预设时间内所述目标子服务集群未接收到所述新的访问请求，则关闭所述目标子服务器集群。

8.一种访问请求的响应装置，其特征在于，包括：

请求接收单元，用于接收用户的访问请求，所述访问请求包含用户标识和用户位置信息；

集群确定单元，用于根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群；

状态获取单元，用于获取所述目标子服务器集群的当前使用状态；

请求数量确定单元，用于确定所述目标子服务集群中所接收的访问请求的数量；

工作参数调整单元，用于根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求；

其中，所述工作参数调整单元在实现所述根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求时，包括：

根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，计算所述目标子服务器集群中当前启动的服务节点的资源使用率；

根据所述目标子服务集群的总用户请求量、总请求处理能力和总请求处理强度，计算所述目标子服务集群的总资源使用率；

若所述资源使用率大于第一阀值且所述总资源使用率大于第三阀值时，则在目标子服务器集群中启动新的服务节点；若所述资源使用率大于所述第一阀值且所述总资源使用率小于第四阀值时，则拒绝接收新的访问请求；若所述资源使用率小于第二阀值时，则在目标子服务器集群中关闭一个当前启动的服务节点，所述第二阀值小于所述第一阀值，所述第四阀值小于所述第三阀值。

9.一种计算机设备，其特征在于，所述计算机设备包括存储器和处理器；

所述存储器用于存储计算机程序；

所述处理器，用于执行所述计算机程序并在执行所述计算机程序时实现如权利要求1至7中任一项所述的访问请求的响应方法。

10.一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有计算机程序，所述计算机程序被处理器执行时使所述处理器实现如权利要求1至7中任一项所述的访问请求的响应方法。

访问请求的响应方法、装置、设备及存储介质

技术领域

[0001] 本申请涉及互联网保险领域，尤其涉及一种访问请求的响应方法、装置、设备及存储介质。

背景技术

[0002] 随着互联网的不断发展，越来越多的企业需要使用集群服务来更好的满足用户的需求。在实际应用中，用户在不同时间段内网络行为的差异性导致访问请求多少差别较大，用户的访问在某些时间段内有大量请求，而在部分时间段内基本没有请求或者仅有少量请求。为了保证系统能够在访问量在最高时仍能够保持较高的可用性，避免系统在访问量最高时出现故障而对业务造成危害的问题，后台服务节点需要一直保持运行。

[0003] 目前，对于大型的应用服务而言，当应用服务启动时，后台服务节点的数量通常是固定的，在服务运行状态下，后台服务器集群响应慢，用户的访问体验不够理想。

发明内容

[0004] 本申请提供了一种访问请求的响应方法、装置、设备及存储介质，该方法能够动态调整服务器集群的服务节点的工作状态，提高服务器集群的响应速度，由此提高用户的访问体验度。

[0005] 第一方面，本申请提供了一种访问请求的响应方法，所述方法包括：

[0006] 接收用户的访问请求，所述访问请求包含用户标识和用户位置信息；

[0007] 根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群；

[0008] 获取所述目标子服务器集群的当前使用状态；

[0009] 确定所述目标子服务集群所接收的访问请求的数量；

[0010] 根据所述接收的访问请求的数量和所述当前使用状态，调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求。

[0011] 第二方面，本申请还提供了一种访问请求的响应装置，所述装置包括：

[0012] 请求接收单元，用于接收用户的访问请求，所述访问请求包含用户标识和用户位置信息；

[0013] 集群确定单元，用于根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群；

[0014] 状态获取单元，用于获取所述目标子服务器集群的当前使用状态；

[0015] 请求数量确定单元，用于确定所述目标子服务集群中所接收的访问请求的数量；

[0016] 工作参数调整单元，用于根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求。

[0017] 第三方面，本申请还提供了一种计算机设备，所述计算机设备包括存储器和处理

器;所述存储器用于存储计算机程序;所述处理器,用于执行所述计算机程序并在执行所述计算机程序时实现如上述的访问请求的响应方法。

[0018] 第四方面,本申请还提供了一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,所述计算机程序被处理器执行时使所述处理器实现如上述的访问请求的响应方法。

[0019] 本申请公开了一种访问请求的响应方法、装置、设备及存储介质,通过接收用户的访问请求;根据所述用户标识和所述用户位置信息,确定响应所述访问请求的目标子服务器集群;获取所述目标子服务器集群的当前使用状态;确定所述目标子服务集群所接收的访问请求的数量;根据所述接收的访问请求的数量和所述当前使用状态,调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求,从而动态调整服务器集群的服务节点的工作状态,提高了服务器集群的响应速度,由此提高了用户的访问体验度,同时也避免了资源浪费。

附图说明

[0020] 为了更清楚地说明本申请实施例技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0021] 图1是本申请的实施例提供的访问请求的响应方法的应用场景示意图;

[0022] 图2是图1中子服务集群的结构示意图;

[0023] 图3是本申请一实施例提供的一种访问请求的响应方法的步骤示意流程图;

[0024] 图4是图3中访问请求的响应方法的子步骤示意流程图;

[0025] 图5是本申请另一实施例提供的一种访问请求的响应方法的步骤示意流程图;

[0026] 图6是本申请又一实施例提供的一种访问请求的响应方法的步骤示意流程图;

[0027] 图7是本申请再一实施例提供的一种访问请求的响应方法的部分步骤示意流程图;

[0028] 图8是本申请的实施例还提供一种访问请求的响应装置的示意性框图;

[0029] 图9是图8中访问请求的响应装置的子单元的示意性框图;

[0030] 图10是图8中访问请求的响应装置的子单元的示意性框图;

[0031] 图11为本申请一实施例提供的一种计算机设备的结构示意性框图。

具体实施方式

[0032] 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,所描述的实施例是本申请一部分实施例,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,都属于本申请保护的范围。

[0033] 附图中所示的流程图仅是示例说明,不是必须包括所有的内容和操作/步骤,也不是必须按所描述的顺序执行。例如,有的操作/步骤还可以分解、组合或部分合并,因此实际执行的顺序有可能根据实际情况改变。

[0034] 本申请的实施例提供了一种访问请求的响应方法、装置、计算机设备及存储介质。

访问请求的响应方法可用于针对患者和/或医生的骗保行为，为快速找出患者或医生骗保提供了重要的参考。

[0035] 下面结合附图，对本申请的一些实施方式作详细说明。在不冲突的情况下，下述的实施例及实施例中的特征可以相互组合。

[0036] 请参阅图1，图1是本申请的实施例提供的访问请求的响应方法的应用场景示意图。本申请的访问请求的响应方法可以应用于由用户终端110和服务器集群120构成的分布式系统中。服务器集群120包括多个子服务集群121。如图2所示，每个子服务集群121包括至少一个服务节点。用户终端110可以有多个，比如用户终端110a和用户终端110b。每个用户终端110可以与所述服务集群120中的任意一个服务节点通信，从而动态调整服务节点的工作参数以响应用户的访问请求。

[0037] 其中，用户终端110可以是各种电子设备，包括但不限于智能手机、平板电脑、膝上便携计算机和台式计算机。

[0038] 具体的，用户终端110上可以安装有各种通讯客户端应用，例如网页浏览器应用、搜索类应用、即时通讯工具、邮箱客户端等。通过用户终端110与服务节点的交互完成对服务节点的工作参数进行动态调整。

[0039] 请参阅图3，图3是本申请一实施例提供的一种访问请求的响应方法的步骤示意流程图。该访问请求的响应方法，用于动态调整服务节点的工作参数。对于某些子服务器集群，用户的访问请求可能是阶段性的，例如，在某个时间段内，访问请求的数量特别多，在某个时间段内没有访问请求。

[0040] 其中，服务器集群是由具有相同功能的多个服务器组成。其中，目标子服务器集群为服务器集群中的至少一个子服务器集群。

[0041] 如图3所示，该访问请求的响应方法，具体包括：步骤S201至步骤S205。

[0042] S201、接收用户的访问请求。

[0043] 具体的，所述访问请求包含用户标识和用户位置信息等信息。在实际应用中，每个用户均具有唯一一个用户标识。用户标识可以包括员工编号或用户姓名等信息。

[0044] 用户位置信息可以包括用户的地理位置信息和/或逻辑网络拓扑位置信息。用户的地理位置信息包括地址标签，和/或，经纬度。地址标签为省、市、县、街道等信息，如北京市、海淀区等、“广东省”，“深圳市”等。

[0045] 用户可以使用该用户终端通过有线连接方式或无线连接方式与目标子服务器集群中当前启动的服务节点交互，以发送访问请求。无线连接方式可以包括但不限于3G/4G连接、WiFi连接、蓝牙连接、WiMAX连接、Zigbee连接、UWB(ultra wideband)连接以及其他目前公知或未来开发的无线连接方式。

[0046] S202、根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群。

[0047] 在本实施例中，步骤S202，具体包括子步骤S202a和S202b。

[0048] 步骤S202a、根据用户标识获得用户的注册地址。

[0049] 步骤S202b、根据用户的注册地址和所述用户位置信息确定响应访问请求的目标子服务器集群。

[0050] 本实施例中，根据上述用户的注册地址和用户的位置信息，来确定响应访问请求

的目标子服务集群，可以保证响应访问请求，可以避免竞争产生脏数据，并且保证每次访问都是离用户最近的子服务器集群，从而提高访问速度，提高用户的体验度。

[0051] S203、获取所述目标子服务器集群的当前使用状态。

[0052] 具体的，所述目标子服务器集群的当前使用状态包括目标子服务器集群中当前启动的服务节点的数量和当前启动的服务节点的请求处理信息。

[0053] S204、确定所述目标子服务集群所接收的访问请求的数量。

[0054] 其中，目标子服务集群中的每一个服务节点可以接收不同用户的访问请求，也可以接收同一用户的不同访问请求。

[0055] 在本实施例中，所述目标子服务集群中各服务节点在预设时间所接收的访问请求的数量之和即为所述目标子服务集群所接收的访问请求的数量。

[0056] 示例性的，目标子服务集群中有两个服务节点，分别为服务节点A和服务节点B，服务节点A在十分钟内接收到的访问请求的数量为2个，服务节点B在十分钟内接收到的访问请求的数量为3个，则目标子服务集群在十分钟内所接收到的访问请求的数量为5个。

[0057] S205、根据所述接收的访问请求的数量和所述当前使用状态，调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求。

[0058] 具体的，所述工作参数包括服务节点每次可响应的访问请求的数量以及服务节点的工作状态。工作状态包括启动状态和关闭状态。

[0059] 在本实施例中，步骤S205之前还包括:判断目标子服务集群中当前是否有启动的服务节点。

[0060] 若目标子服务集群中当前没有启动的服务节点，则启动目标子服务集群。若目标子服务集群中当前有启动的服务节点，则执行步骤S205。如果目标子服务集群的当前使用状态指示目标子服务集群中当前启动的服务节点的数量为零，则表明目标子服务集群当前没有启动的服务节点。

[0061] 其中，所述根据所述接收的访问请求的数量和所述当前使用状态，调整所述目标子服务器集群中服务节点的工作参数，具体为：

[0062] 根据所述接收的访问请求的数量、所述目标子服务集群中已启动的服务节点的请求处理能力和所述目标子服务集群中已启动的服务节点的数量，调整所述目标子服务器集群中服务节点的工作参数。

[0063] 示例性的，目标子服务集群有8个服务节点，其中3个服务节点的工作状态为启动状态，剩余5个服务节点的工作状态为关闭状态。若当前用户的访问请求的数量较大，3个服务节点的服务能力无法满足当前用户的访问请求时，则调整剩余5个服务节点的工作状态。需要调整工作状态的服务节点的数量依据当前用户的访问请求的数量和当前已启动的服务节点的资源使用情况确定，比如，需要再启动5个服务节点才能满足用户的访问请求，则将剩余的5个服务节点的工作状态由关闭状态调整为启动状态。

[0064] 上述访问请求的响应方法，通过接收用户的访问请求；根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群；获取所述目标子服务器集群的当前使用状态；确定所述目标子服务集群所接收的访问请求的数量；根据所述接收的访问请求的数量和所述当前使用状态，调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求，从而动态调整服务器集群的服务节点的工作状态，提高了服务

器集群的响应速度，由此提高了用户的访问体验度，同时也避免了资源浪费。

[0065] 请参阅图5，图5是本申请另一实施例提供的一种访问请求的响应方法的步骤示意流程图。该访问请求的响应方法，用于动态调整服务节点的工作参数。对于某些子服务器集群，用户的访问请求可能是阶段性的，例如，在某个时间段内，访问请求的数量特别多，在某个时间段内没有访问请求。

[0066] 其中，服务器集群是由具有相同功能的多个服务器组成。其中，目标子服务器集群为服务器集群中的至少一个子服务器集群。

[0067] 如图5所示，该访问请求的响应方法，具体包括：步骤S301至步骤S306。

[0068] S301、接收用户的访问请求。

[0069] 具体的，所述访问请求包含用户标识和用户位置信息等信息。在实际应用中，每个用户均具有唯一一个用户标识。用户标识可以包括员工编号或用户姓名等信息。

[0070] 用户位置信息可以包括用户的地理位置信息和/或逻辑网络拓扑位置信息。用户的地理位置信息包括地址标签，和/或，经纬度。地址标签为省、市、县、街道等信息，如北京市、海淀区等、“广东省”，“深圳市”等。

[0071] S302、根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群。

[0072] 在本实施例中，步骤S302具体包括：根据用户标识获得用户的注册地址；根据用户的注册地址和用户的位置信息确定响应访问请求的目标子服务器集群，可以避免竞争产生脏数据，并且保证每次访问都是离用户最近的子服务器集群，从而提高访问速度，提高用户的体验度。

[0073] S303、获取所述目标子服务器集群的当前使用状态。

[0074] 所述目标子服务器集群的当前使用状态包括目标子服务器集群中当前启动的服务节点的数量和当前启动的服务节点的请求处理信息。

[0075] 请求处理信息包括请求处理能力、最大处理能力、请求处理强度、最大处理强度、请求处理强度、最大处理强度以及当前启动的服务节点的访问请求的并行处理能力。

[0076] S304、确定所述目标子服务集群所接收的访问请求的数量。

[0077] 具体的，目标子服务集群中的每一个服务节点可以接收不同用户的访问请求，也可以接收同一用户的不同访问请求。

[0078] 在本实施例中，所述目标子服务集群中各服务节点在预设时间所接收的访问请求的数量之和即为所述目标子服务集群所接收的访问请求的数量。

[0079] S305、基于使用率计算公式，根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，计算所述目标子服务器集群中当前启动的服务节点的资源使用率。

[0080] 在本实施例中，所述使用率计算公式为：

[0081] $$T_i=\begin{cases}1, & if\,(r_i=R_i\ or\ q_i=Q_i\ or\ s_i=S_i)\\ w_1\dfrac{r_i}{R_i}+w_2\dfrac{q_i}{Q_i}+w_3\dfrac{s_i}{S_i}, & Others\end{cases}$$

[0082] 其中，T_i表示目标子服务集群中当前启动的第i个服务节点的资源使用率，r_i、R_i分别表示服务节点的用户请求量和最大请求量，q_i、Q_i分别表示服务节点的请求处理能力和最

大处理能力，s_i、S_i分别表示服务节点的请求处理强度和最大处理强度，$s_i=\frac{r_i}{C_i \cdot q_i}$，$C_i$为目标子服务集群中当前启动的第i个服务节点的访问请求的并行处理能力，w_1、w_2、w_3分别表示用户请求量、请求处理能力和请求处理强度的权重，且$w_1+w_2+w_3=1$。

[0083] 具体的，假设目标子服务器集群Y中具有N个当前启动的服务节点，即$Y=\{Y_1, Y_2, \ldots, Y_n\}$。用户请求量$r_i$指第i个服务节点$Y_i$在单位时间内收到的当前用户的平均访问请求的数量。请求处理能力q_i指第i个服务节点Y_i在单位时间内完成的当前用户的平均访问请求的数量，其动态显示了当前用户的访问请求的处理能力。请求处理强度s_i体现了服务节点Y_i的负载状况与实际请求处理能力的相对关系。请求处理强度s_i越接近1说明服务节点Y_i的负载状态越大。

[0084] 基于上述使用率计算公式，根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，可以计算出所述目标子服务器集群中当前启动的服务节点的资源使用率。

[0085] S306、若所述资源使用率大于第一阀值时，则在所述目标子服务器集群中启动新的服务节点；若所述资源使用率小于第二阀值时，则在所述目标子服务器集群中关闭一个当前启动的服务节点。

[0086] 其中，服务节点Y_i的负载状态存在过载、空闲、正常三种状态，所述第二阀值α_2小于所述第一阀值α_1。当资源使用率T_i大于第一阀值α_1时，服务节点Y_i的负载状态为过载状态，此时，目标子服务器集群中启动的服务节点的数量小于当前用户请求的服务所需要的服务节点的数量。当资源使用率T_i小于第二阀值α_2($\alpha_2<\alpha_1$)时，服务节点Y_i的负载状态为空闲状态，此时，目标子服务器集群中启动的服务节点的数量大于当前用户请求的服务所需要的服务节点的数量。当$\alpha_2 \leqslant T_i \leqslant \alpha_1$时，服务节点$Y_i$的负载状态为正常状态。当目标子服务器集群当前没有启动的服务节点时，根据访问请求的数量调整目标子服务器集群中服务节点的工作参数以响应访问请求。

[0087] 其中，第一阀值α_1和第二阀值α_2可以根据实际需求进行设计。例如，第一阀值α_1取值为85％，第二阀值α_2取值为5％，若资源使用率T_i为90％，则在所述目标子服务集群中启动新的服务节点，以使当前已启动的服务节点的服务能力满足用户的访问请求。若资源使用率T_i为2％，则表示存在部分空闲的服务节点，此时，在目标子服务器集群中关闭当前启动的服务节点Y_i，当然也可以关闭所述目标子服务集群中其他任意一个当前已启动的服务节点，以提高目标子服务器集群的使用率，大大降低服务器集群的功耗。若资源使用率T_i为60％，则保持所述目标子服务集群中各服务节点的当前工作状态，即既不启动新的服务节点、也不关闭当前已启动的服务节点。

[0088] 当资源使用率$T_i=0$时，表示目标子服务器集群当前没有启动的服务节点，此时，启动目标子服务器集群。具体的，根据访问请求的数量确定目标子服务器集群中需要启动的服务节点的数量并启动。

[0089] 上述实施例提供的访问请求的响应方法，接收用户的访问请求；根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群；获取所述目标子服务器集群的当前使用状态；确定所述目标子服务集群所接收的访问请求的数量；基于使用率计算公式，根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，

计算所述目标子服务器集群中当前启动的服务节点的资源使用率;若所述资源使用率大于第一阀值时,则在所述目标子服务器集群中启动新的服务节点;若所述资源使用率小于第二阀值时,则在所述目标子服务器集群中关闭一个当前启动的服务节点,从而动态调整服务器集群的服务节点的工作状态,提高了服务器集群的响应速度,由此提高了用户的访问体验度,同时也避免了资源浪费。

[0090] 请参阅图6,图6是本申请又一实施例提供的一种访问请求的响应方法的步骤示意流程图。该访问请求的响应方法,用于动态调整服务节点的工作参数。目标子服务器集群为服务器集群中的至少一个子服务器集群。

[0091] 如图6所示,该访问请求的响应方法,具体包括:步骤S401至步骤S407。

[0092] S401、接收用户的访问请求。

[0093] 具体的,所述访问请求包含用户标识和用户位置信息等信息。在实际应用中,每个用户均具有唯一一个用户标识。用户标识可以包括员工编号或用户姓名等信息。

[0094] 用户位置信息可以包括用户的地理位置信息和/或逻辑网络拓扑位置信息。用户的地理位置信息包括地址标签,和/或,经纬度。地址标签为省、市、县、街道等信息,如北京市、海淀区等、“广东省”,“深圳市”等。

[0095] S402、根据所述用户标识和所述用户位置信息,确定响应所述访问请求的目标子服务器集群。

[0096] 在本实施例中,步骤S402,具体包括:根据用户标识获得用户的注册地址;根据用户的注册地址和用户位置信息确定响应访问请求的目标子服务器集群,可以避免竞争产生脏数据,并且保证每次访问都是离用户最近的子服务器集群,从而提高访问速度,提高用户的体验度。

[0097] S403、获取所述目标子服务器集群的当前使用状态。

[0098] 所述目标子服务器集群的当前使用状态包括目标子服务器集群中当前启动的服务节点的数量和当前启动的服务节点的请求处理信息。

[0099] 请求处理信息包括请求处理能力、最大处理能力、请求处理强度、最大处理强度、请求处理强度、最大处理强度以及当前启动的服务节点的访问请求的并行处理能力。

[0100] S404、确定所述目标子服务集群所接收的访问请求的数量。

[0101] 具体的,目标子服务集群中的每一个服务节点可以接收不同用户的访问请求,也可以接收同一用户的不同访问请求。

[0102] 在本实施例中,所述目标子服务集群中各服务节点在预设时间所接收的访问请求的数量之和即为所述目标子服务集群所接收的访问请求的数量。

[0103] S405、基于使用率计算公式,根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态,计算所述目标子服务器集群中当前启动的服务节点的资源使用率。

[0104] 在本实施例中,所述使用率计算公式为:

[0105] $$T_i=\begin{cases}1, & if\,(r_i=R_i\ \ or\ \ q_i=Q_i\ \ or\ \ s_i=S_i)\\ w_1\dfrac{r_i}{R_i}+w_2\dfrac{q_i}{Q_i}+w_3\dfrac{s_i}{S_i}, & Others\end{cases}$$

[0106] 其中,T_i表示目标子服务集群中当前启动的第i个服务节点的资源使用率,r_i、R_i分

别表示服务节点的用户请求量和最大请求量，q_i、Q_i分别表示服务节点的请求处理能力和最大处理能力，s_i、S_i分别表示服务节点的请求处理强度和最大处理强度，$s_i=\frac{r_i}{C_i\cdot q_i}$，$C_i$为目标子服务集群中当前启动的第i个服务节点的访问请求的并行处理能力，w_1、w_2、w_3分别表示用户请求量、请求处理能力和请求处理强度的权重，且$w_1+w_2+w_3=1$。

[0107] S406、基于总使用率计算公式，根据所述目标子服务集群的总用户请求量、总请求处理能力和总请求处理强度，计算所述目标子服务集群的总资源使用率。

[0108] 其中，所述总使用率计算公式为：

[0109] $$T_N=\begin{cases}1, & if(r_N=R_N\ or\ q_N=Q_N\ or\ s_N=S_N)\\ w_4\frac{r_N}{R_N}+w_5\frac{q_N}{Q_N}+w_6\frac{s_N}{S_N}, & Others\end{cases}$$

[0110] 其中，T_N表示目标子服务集群的总资源使用率，r_N、R_N分别表示目标子服务集群的总用户请求量和最大总请求量，q_N、Q_N分别表示目标子服务集群的总请求处理能力和最大总处理能力，s_N、S_N分别表示目标子服务集群的总请求处理强度和最大总请求处理强度，w_4，w_5，w_6分别表示总用户请求量、总请求处理能力和总请求处理强度的权重，且$w_4+w_5+w_6=1$。

[0111] S407、若所述资源使用率大于所述第一阀值且所述总资源使用率大于第三阀值时，则在目标子服务器集群中启动新的服务节点；若所述资源使用率大于所述第一阀值且所述总资源使用率小于第四阀值时，则拒绝接收新的访问请求；若所述资源使用率小于所述第二阀值时，则在目标子服务器集群中关闭一个当前启动的服务节点。

[0112] 其中，所述第四阀值α_4小于所述第三阀值α_3。示例性的，第一阀值α_1取值为85％，第二阀值α_2取值为5％，第三阀值α_3取值为75％，第四阀值α_4取值为10％。若资源使用率T_i为90％、总资源使用率T_N为80％，则在所述目标子服务集群中启动新的服务节点，以使当前已启动的服务节点的服务能力满足用户的访问请求。

[0113] 若资源使用率T为90％、总资源使用率T_N为6％，则表示已启动的服务节点中存在资源使用率小于90％的服务节点，即存在部分空闲的服务节点，此时，资源使用率T_i为90％的服务节点拒绝接收新的访问请求，资源使用率T_i小于90％的服务节点可以接收新的访问请求。

[0114] 若资源使用率T_i为2％，则表示该服务节点处于空闲状态，此时，关闭该当前启动的服务节点Y_i，当然也可以关闭所述目标子服务集群中其他任意一个当前已启动的服务节点，以提高目标子服务器集群的使用率，大大降低服务器集群的功耗。

[0115] 对于某些子服务器集群，用户的访问请求可能是阶段性的，例如，在某个时间段内，访问请求的数量特别多，在某个时间段内没有访问请求。在本实施例中，为了进一步提高整个服务集群的利用率、降低功耗，步骤307之后还包括：判断在预设时间内所述目标子服务集群是否接收到新的访问请求。若在预设时间内所述目标子服务集群未接收到所述新的访问请求，则关闭所述目标子服务器集群。

[0116] 在本实施中，在步骤S406之前，还包括：

[0117] S501、基于总请求量计算公式，根据所述目标子服务集群的当前启动的各服务节点的用户请求量，计算所述目标子服务集群的总用户请求量。

[0118] 在本实施例中，所述总请求量计算公式为：

[0119] $$r_N=\sum_{i=1}^{n}r_i$$

[0120] 其中，r_N表示目标子服务集群的总用户请求量，r_i表示目标子服务集群的当前启动的第i个服务节点的用户请求量，n为目标子服务集群中当前启动的服务节点的数量。

[0121] S502、基于总处理能力计算公式，根据所述目标子服务集群的当前启动的各服务节点的请求处理能力，计算所述目标子服务集群的总请求处理能力。

[0122] 在本实施例中，所述总处理能力计算公式为：

[0123] $$q_N=n/\sum_{i=1}^{n}\frac{1}{q_i}$$

[0124] 其中，q_N表示目标子服务集群的总请求处理能力，q_i表示目标子服务集群的当前启动的第i个服务节点的请求处理能力。

[0125] S503、基于总处理强度计算公式，根据所述总用户请求量和总请求处理能力，计算所述目标子服务集群的总请求处理强度。

[0126] 在本实施例中，所述总处理强度计算公式为：

[0127] $$s_N=\frac{r_N}{(q_N\cdot\sum_{i=1}^{n}C_i)}$$

[0128] 其中，s_N表示目标子服务集群的总请求处理强度。

[0129] 上述实施例提供的访问请求的响应方法，接收用户的访问请求；根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群；获取所述目标子服务器集群的当前使用状态；确定所述目标子服务集群所接收的访问请求的数量；基于使用率计算公式，根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，计算所述目标子服务器集群中当前启动的服务节点的资源使用率；基于总使用率计算公式，根据所述目标子服务集群的总用户请求量、总请求处理能力和总请求处理强度，计算所述目标子服务集群的总资源使用率；若所述资源使用率大于所述第一阀值且所述总资源使用率大于第三阀值时，则在目标子服务器集群中启动新的服务节点；若所述资源使用率大于所述第一阀值且所述总资源使用率小于第四阀值时，则拒绝接收新的访问请求；若所述资源使用率小于所述第二阀值时，则在目标子服务器集群中关闭一个当前启动的服务节点，从而动态调整服务器集群的服务节点的工作状态，提高了服务器集群的响应速度，由此提高了用户的访问体验度，同时也避免了资源浪费。

[0130] 请参阅图8，图8是本申请的实施例还提供一种访问请求的响应装置的示意性框图，该访问请求的响应装置用于执行前述任一项访问请求的响应方法。其中，该访问请求的响应装置可以配置于服务器或终端中。

[0131] 其中，服务器可以为独立的服务器，也可以为服务器集群。该终端可以是手机、平板电脑、笔记本电脑、台式电脑、个人数字助理和穿戴式设备等电子设备。

[0132] 如图8所示，访问请求的响应装置600包括：

[0133] 请求接收单元601，用于接收用户的访问请求，所述访问请求包含用户标识和用户位置信息；

[0134] 集群确定单元602，用于根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群；

[0135] 状态获取单元603，用于获取所述目标子服务器集群的当前使用状态；

[0136] 请求数量确定单元604，用于确定所述目标子服务集群所接收的访问请求的数量；

[0137] 工作参数调整单元605，用于根据所述接收的访问请求的数量和所述当前使用状态，调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求。

[0138] 在一个实施例中，如图9所示，集群确定单元602包括地址获取子单元602a和集群确定子单元602b。

[0139] 地址获取子单元602a，用于根据所述用户标识获取用户的注册地址。

[0140] 集群确定子单元602b，用于根据所述注册地址和所述用户位置信息确定响应所述访问请求的目标子服务集群。

[0141] 在一个实施例中，如图10所示，工作参数调整单元605包括使用率计算子单元605a和启动关闭单元605b。

[0142] 使用率计算子单元605a，用于基于使用率计算公式，根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，计算所述目标子服务器集群中当前启动的服务节点的资源使用率；

[0143] 启动关闭单元605b，用于若所述资源使用率大于第一阀值时，则在所述目标子服务器集群中启动新的服务节点；若所述资源使用率小于第二阀值时，则在所述目标子服务器集群中关闭一个当前启动的服务节点；其中，所述第二阀值小于所述第一阀值。

[0144] 在一实施例中，所述使用率计算公式为：

[0145] $$T_i=\begin{cases}1, & if\,(r_i=R_i\ \ or\ \ q_i=Q_i\ \ or\ \ s_i=S_i)\\ w_1\dfrac{r_i}{R_i}+w_2\dfrac{q_i}{Q_i}+w_3\dfrac{s_i}{S_i}, & Others\end{cases}$$

[0146] 其中，T_i表示目标子服务集群中当前启动的第i个服务节点的资源使用率，r_i、R_i分别表示服务节点的用户请求量和最大请求量，q_i、Q_i分别表示服务节点的请求处理能力和最大处理能力，s_i、S_i分别表示服务节点的请求处理强度和最大处理强度，$S_i=\dfrac{r_i}{C_i\cdot q_i}$，$C_i$为目标子服务集群中当前启动的第i个服务节点的请求的并行处理能力，w_1、w_2、w_3分别表示用户请求量、请求处理能力和请求处理强度的权重，且$w_1+w_2+w_3=1$。

[0147] 在一实施例中，工作参数调整单元605还包括总使用率计算子单元605c。

[0148] 总使用率计算子单元605c，用于基于总使用率计算公式，根据所述目标子服务集群的总用户请求量、总请求处理能力和总请求处理强度，计算所述目标子服务集群的总资源使用率。

[0149] 在该实施例中，启动关闭单元605b，用于若所述资源使用率大于所述第一阀值且所述总资源使用率大于第三阀值时，则在目标子服务器集群中启动新的服务节点；若所述资源使用率大于所述第一阀值且所述总资源使用率小于第四阀值时，则拒绝接收新的访问请求；若所述资源使用率小于所述第二阀值时，则在目标子服务器集群中关闭一个当前启动的服务节点；其中，所述第四阀值小于所述第三阀值。

[0150] 具体的，所述总使用率计算公式为：

[0151] $$T_N=\begin{cases}1, & if(r_N=R_N \ or \ q_N=Q_N \ or \ s_N=S_N)\\ w_4\frac{r_N}{R_N}+w_5\frac{q_N}{Q_N}+w_6\frac{s_N}{S_N}, & Others\end{cases}$$

[0152] 其中，T_N表示目标子服务集群的总资源使用率，r_N、R_N分别表示目标子服务集群的总用户请求量和最大总请求量，q_N、Q_N分别表示目标子服务集群的总请求处理能力和最大总处理能力，s_N、S_N分别表示目标子服务集群的总请求处理强度和最大总请求处理强度，w_4，w_5，w_6分别表示总用户请求量、总请求处理能力和总请求处理强度的权重，且$w_4+w_5+w_6=1$。

[0153] 在一实施例中，访问请求的响应装置还包括总用户请求量计算单元、总请求处理能力计算单元和总请求处理强度计算单元。

[0154] 总用户请求量计算单元，用于基于总请求量计算公式，根据所述目标子服务集群的当前启动的各服务节点的用户请求量，计算所述目标子服务集群的总用户请求量。

[0155] 总请求处理能力计算单元，用于基于总处理能力计算公式，根据所述目标子服务集群的当前启动的各服务节点的请求处理能力，计算所述目标子服务集群的总请求处理能力；

[0156] 总请求处理强度计算单元，用于基于总处理强度计算公式，根据所述总用户请求量和总请求处理能力，计算所述目标子服务集群的总请求处理强度。

[0157] 具体的，所述总请求量计算公式为：

[0158] $$r_N=\sum_{i=1}^{n}r_i$$

[0159] 其中，r_N表示目标子服务集群的总用户请求量，r_i表示目标子服务集群的当前启动的第i个服务节点的用户请求量，n为目标子服务集群中当前启动的服务节点的数量。

[0160] 所述总处理能力计算公式为：

[0161] $$q_N=n/\sum_{i=1}^{n}\frac{1}{q_i}$$

[0162] 其中，q_N表示目标子服务集群的总请求处理能力，q_i表示目标子服务集群的当前启动的第i个服务节点的请求处理能力；

[0163] 所述总处理强度计算公式为：

[0164] $$s_N=\frac{r_N}{(q_N\cdot\sum_{i=1}^{n}C_i)}$$

[0165] 其中，s_N表示目标子服务集群的总请求处理强度，C_i为目标子服务集群中当前启动的第i个服务节点的请求的并行处理能力。

[0166] 需要说明的是，所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，上述描述的访问请求的响应装置和各单元的具体工作过程，可以参考前述访问请求的响应方法实施例中的对应过程，在此不再赘述。

[0167] 上述的访问请求的响应装置可以实现为一种计算机程序的形式，该计算机程序可以在如图11所示的计算机设备上运行。

[0168] 请参阅图11,图11是本申请实施例提供的一种计算机设备的示意性框图。该计算机设备可以是服务器或终端。

[0169] 参阅图11,该计算机设备包括通过系统总线连接的处理器、存储器和网络接口,其中,存储器可以包括非易失性存储介质和内存储器。

[0170] 非易失性存储介质可存储操作系统和计算机程序。该计算机程序包括程序指令,该程序指令被执行时,可使得处理器执行一种访问请求的响应方法。

[0171] 处理器用于提供计算和控制能力,支撑整个计算机设备的运行。

[0172] 内存储器为非易失性存储介质中的计算机程序的运行提供环境,该计算机程序被处理器执行时,可使得处理器执行一种访问请求的响应方法。

[0173] 该网络接口用于进行网络通信,如发送分配的任务等。本领域技术人员可以理解,图11中示出的结构,仅仅是与本申请方案相关的部分结构的框图,并不构成对本申请方案所应用于其上的计算机设备的限定,具体的计算机设备可以包括比图中所示更多或更少的部件,或者组合某些部件,或者具有不同的部件布置。

[0174] 应当理解的是,处理器可以是中央处理单元(Central Processing Unit,CPU),该处理器还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现场可编程门阵列(Field-Programmable GateArray,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。其中,通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0175] 其中,所述处理器用于运行存储在存储器中的计算机程序,以实现如下步骤:

[0176] 接收用户的访问请求,所述访问请求包含用户标识和用户位置信息;

[0177] 根据所述用户标识和所述用户位置信息,确定响应所述访问请求的目标子服务器集群;

[0178] 获取所述目标子服务器集群的当前使用状态;

[0179] 确定所述目标子服务集群所接收的访问请求的数量;

[0180] 根据所述接收的访问请求的数量和所述当前使用状态,调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求。

[0181] 在一个实施例中,所述处理器在实现所述根据所述用户标识和所述用户位置信息,确定响应所述访问请求的目标子服务器集群时,用于实现:

[0182] 根据所述用户标识获取用户的注册地址;

[0183] 根据所述注册地址和所述用户位置信息确定响应所述访问请求的目标子服务集群。

[0184] 在一个实施例中,所述处理器在实现所述根据所述接收的访问请求的数量和所述当前使用状态,调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求时,用于实现:

[0185] 基于使用率计算公式,根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态,计算所述目标子服务器集群中当前启动的服务节点的资源使用率;

[0186] 若所述资源使用率大于第一阀值时,则在所述目标子服务器集群中启动新的服务节点;若所述资源使用率小于第二阀值时,则在所述目标子服务器集群中关闭一个当前启

动的服务节点;其中,所述第二阀值小于所述第一阀值。

[0187] 在一个实施例中,所述使用率计算公式为:

[0188] $$T_i=\begin{cases}1, & if(r_i=R_i\ or\ q_i=Q_i\ or\ s_i=S_i)\\ w_1\frac{r_i}{R_i}+w_2\frac{q_i}{Q_i}+w_3\frac{s_i}{S_i}, & Others\end{cases}$$

[0189] 其中,T_i表示目标子服务集群中当前启动的第i个服务节点的资源使用率,r_i、R_i分别表示服务节点的用户请求量和最大请求量,q_i、Q_i分别表示服务节点的请求处理能力和最大处理能力,s_i、S_i分别表示服务节点的请求处理强度和最大处理强度,$S_i=\frac{r_i}{C_i\cdot q_i}$,$C_i$为目标子服务集群中当前启动的第i个服务节点的访问请求的并行处理能力,w_1、w_2、w_3分别表示用户请求量、请求处理能力和请求处理强度的权重,且$w_1+w_2+w_3=1$。

[0190] 在一个实施例中,所述处理器在实现所述基于使用率计算公式,根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态,计算所述目标子服务器集群中当前启动的服务节点的资源使用率时,用于实现:

[0191] 基于总使用率计算公式,根据所述目标子服务集群的总用户请求量、总请求处理能力和总请求处理强度,计算所述目标子服务集群的总资源使用率;

[0192] 若所述资源使用率大于所述第一阀值且所述总资源使用率大于第三阀值时,则在目标子服务器集群中启动新的服务节点;若所述资源使用率大于所述第一阀值且所述总资源使用率小于第四阀值时,则拒绝接收新的访问请求;若所述资源使用率小于所述第二阀值时,则在目标子服务器集群中关闭一个当前启动的服务节点;其中,所述第四阀值小于所述第三阀值;

[0193] 所述总使用率计算公式为:

[0194] $$T_N=\begin{cases}1, & if(r_N=R_N\ or\ q_N=Q_N\ or\ s_N=S_N)\\ w_4\frac{r_N}{R_N}+w_5\frac{q_N}{Q_N}+w_6\frac{s_N}{S_N}, & Others\end{cases}$$

[0195] 其中,T_N表示目标子服务集群的总资源使用率,r_N、R_N分别表示目标子服务集群的总用户请求量和最大总请求量,q_N、Q_N分别表示目标子服务集群的总请求处理能力和最大总处理能力,s_N、S_N分别表示目标子服务集群的总请求处理强度和最大总请求处理强度,w_4,w_5,w_6分别表示总用户请求量、总请求处理能力和总请求处理强度的权重,且$w_4+w_5+w_6=1$。

[0196] 在一实施例中,所述处理器在实现所述基于总使用率计算公式,根据所述目标子服务集群的总用户请求量、总请求处理能力和总请求处理强度,计算所述目标子服务集群的总资源使用率之前,还用于实现:

[0197] 基于总请求量计算公式,根据所述目标子服务集群的当前启动的各服务节点的用户请求量,计算所述目标子服务集群的总用户请求量;

[0198] 基于总处理能力计算公式,根据所述目标子服务集群的当前启动的各服务节点的请求处理能力,计算所述目标子服务集群的总请求处理能力;

[0199] 基于总处理强度计算公式,根据所述总用户请求量和总请求处理能力,计算所述目标子服务集群的总请求处理强度;

[0200] 所述总请求量计算公式为：

[0201] $$r_N=\sum_{i=1}^{n}r_i$$

[0202] 其中，r_N表示目标子服务集群的总用户请求量，r_i表示目标子服务集群的当前启动的第i个服务节点的用户请求量，n为目标子服务集群中当前启动的服务节点的数量；

[0203] 所述总处理能力计算公式为：

[0204] $$q_N=n/\sum_{i=1}^{n}\frac{1}{q_i}$$

[0205] 其中，q_N表示目标子服务集群的总请求处理能力，q_i表示目标子服务集群的当前启动的第i个服务节点的请求处理能力；

[0206] 所述总处理强度计算公式为：

[0207] $$s_N=\frac{r_N}{(q_N\cdot\sum_{i=1}^{n}C_i)}$$

[0208] 其中，s_N表示目标子服务集群的总请求处理强度，C_i为目标子服务集群中当前启动的第i个服务节点的访问请求的并行处理能力。

[0209] 本申请的实施例中还提供一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序中包括程序指令，所述处理器执行所述程序指令，实现本申请实施例提供的任一项访问请求的响应方法。

[0210] 其中，所述计算机可读存储介质可以是前述实施例所述的计算机设备的内部存储单元，例如所述计算机设备的硬盘或内存。所述计算机可读存储介质也可以是所述计算机设备的外部存储设备，例如所述计算机设备上配备的插接式硬盘，智能存储卡(Smart Media Card，SMC)，安全数字(Secure Digital，SD)卡，闪存卡(Flash Card)等。

[0211] 以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到各种等效的修改或替换，这些修改或替换都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以权利要求的保护范围为准。

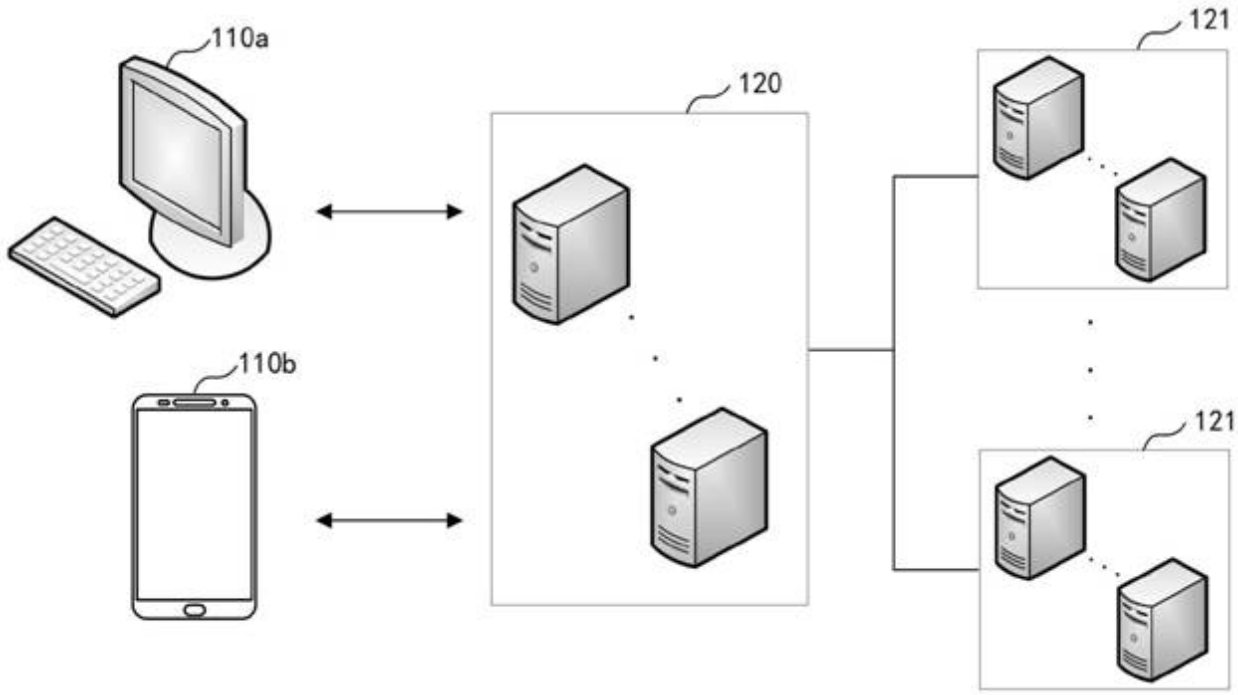


图1

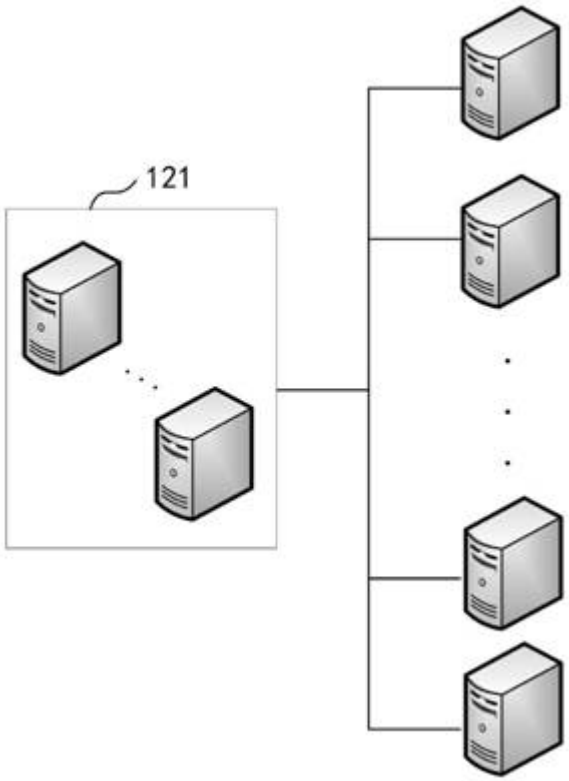


图2

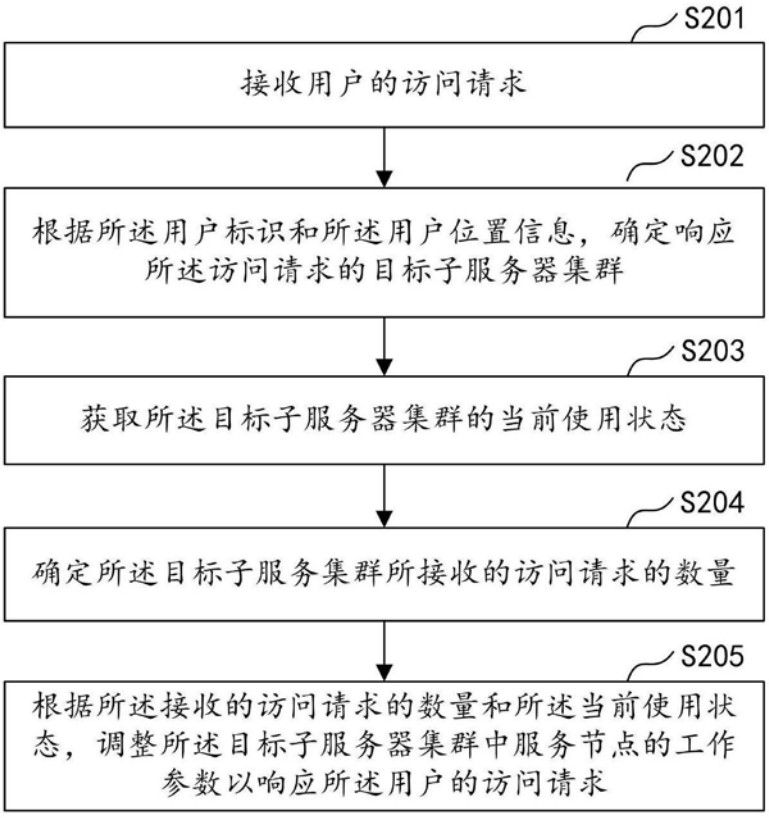

S201
接收用户的访问请求
S202
根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群
S203
获取所述目标子服务器集群的当前使用状态
S204
确定所述目标子服务集群所接收的访问请求的数量
S205
根据所述接收的访问请求的数量和所述当前使用状态，调整所述目标子服务器集群中服务节点的工作参数以响应所述用户的访问请求

图3

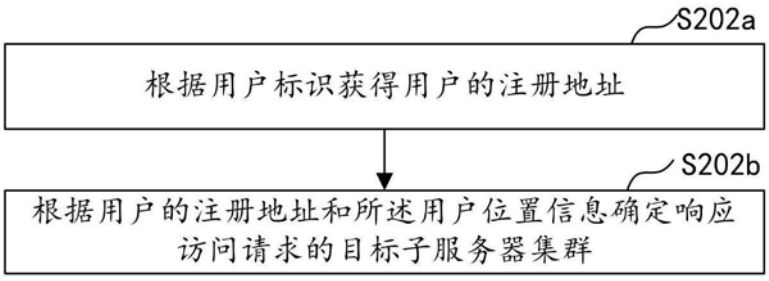

S202a
根据用户标识获得用户的注册地址
S202b
根据用户的注册地址和所述用户位置信息确定响应访问请求的目标子服务器集群

图4

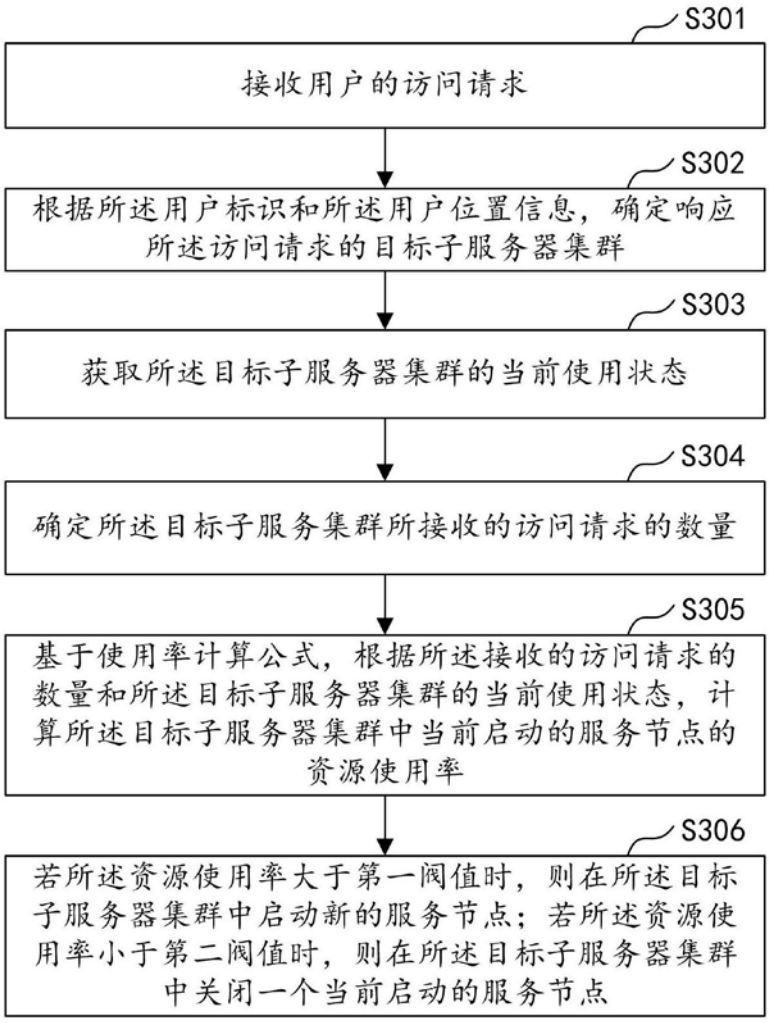

S301
接收用户的访问请求
S302
根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群
S303
获取所述目标子服务器集群的当前使用状态
S304
确定所述目标子服务集群所接收的访问请求的数量
S305
基于使用率计算公式，根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，计算所述目标子服务器集群中当前启动的服务节点的资源使用率
S306
若所述资源使用率大于第一阈值时，则在所述目标子服务器集群中启动新的服务节点；若所述资源使用率小于第二阈值时，则在所述目标子服务器集群中关闭一个当前启动的服务节点

图5

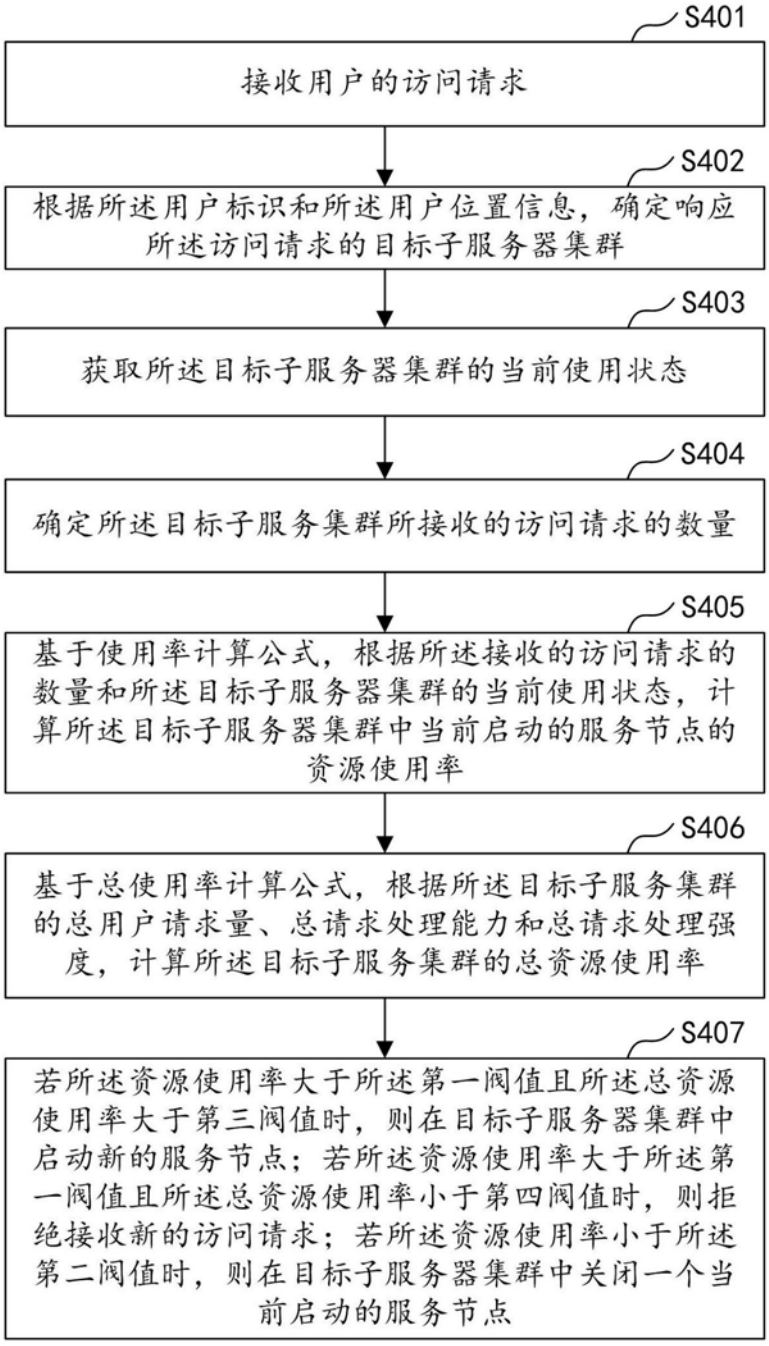

S401
接收用户的访问请求
S402
根据所述用户标识和所述用户位置信息，确定响应所述访问请求的目标子服务器集群
S403
获取所述目标子服务器集群的当前使用状态
S404
确定所述目标子服务集群所接收的访问请求的数量
S405
基于使用率计算公式，根据所述接收的访问请求的数量和所述目标子服务器集群的当前使用状态，计算所述目标子服务器集群中当前启动的服务节点的资源使用率
S406
基于总使用率计算公式，根据所述目标子服务集群的总用户请求量、总请求处理能力和总请求处理强度，计算所述目标子服务集群的总资源使用率
S407
若所述资源使用率大于所述第一阈值且所述总资源使用率大于第三阈值时，则在目标子服务器集群中启动新的服务节点；若所述资源使用率大于所述第一阈值且所述总资源使用率小于第四阈值时，则拒绝接收新的访问请求；若所述资源使用率小于所述第二阈值时，则在目标子服务器集群中关闭一个当前启动的服务节点

图6

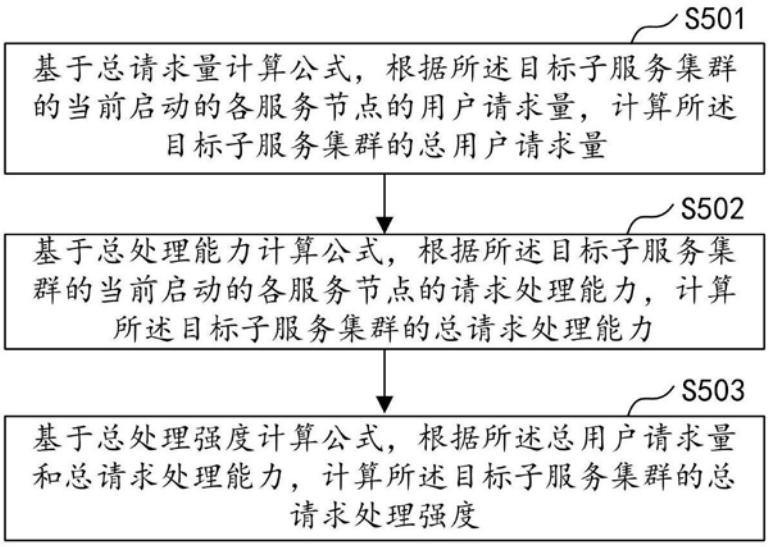

S501
基于总请求量计算公式，根据所述目标子服务集群的当前启动的各服务节点的用户请求量，计算所述目标子服务集群的总用户请求量
S502
基于总处理能力计算公式，根据所述目标子服务集群的当前启动的各服务节点的请求处理能力，计算所述目标子服务集群的总请求处理能力
S503
基于总处理强度计算公式，根据所述总用户请求量和总请求处理能力，计算所述目标子服务集群的总请求处理强度

图7

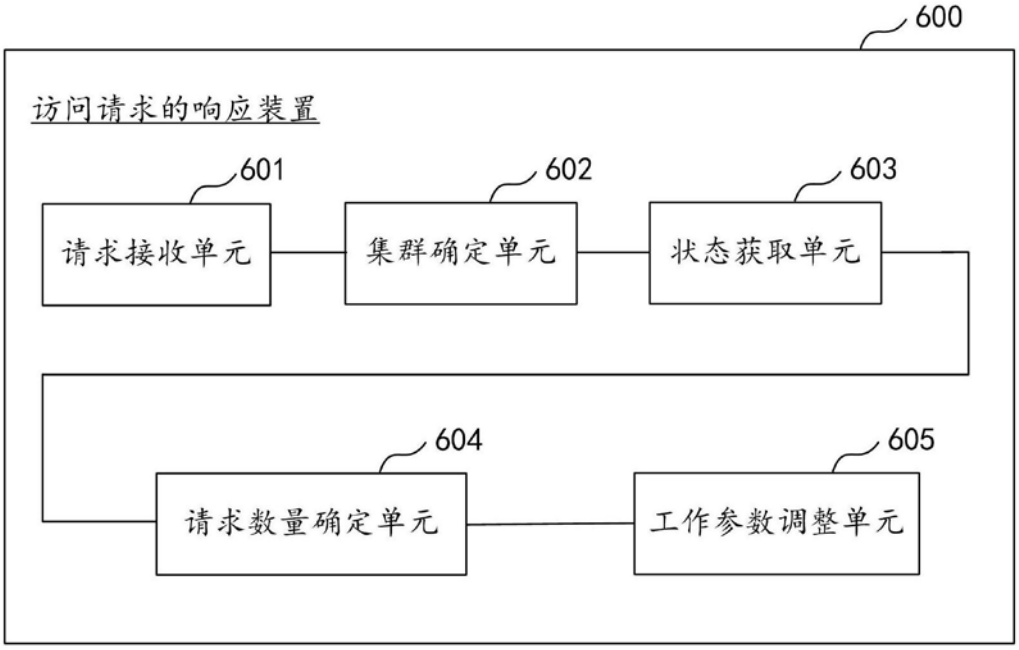

600
访问请求的响应装置
601
请求接收单元
602
集群确定单元
603
状态获取单元
604
请求数量确定单元
605
工作参数调整单元

图8

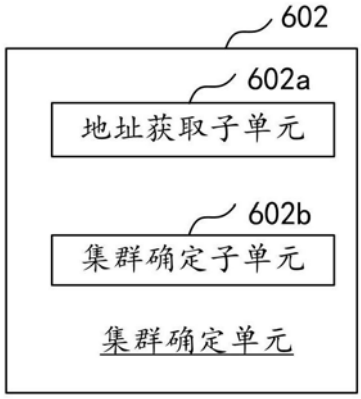

602
602a
地址获取子单元
602b
集群确定子单元
集群确定单元

图9

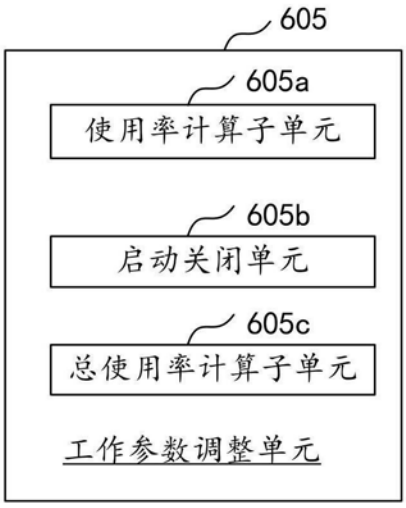

605
605a
使用率计算子单元
605b
启动关闭单元
605c
总使用率计算子单元
工作参数调整单元

图10

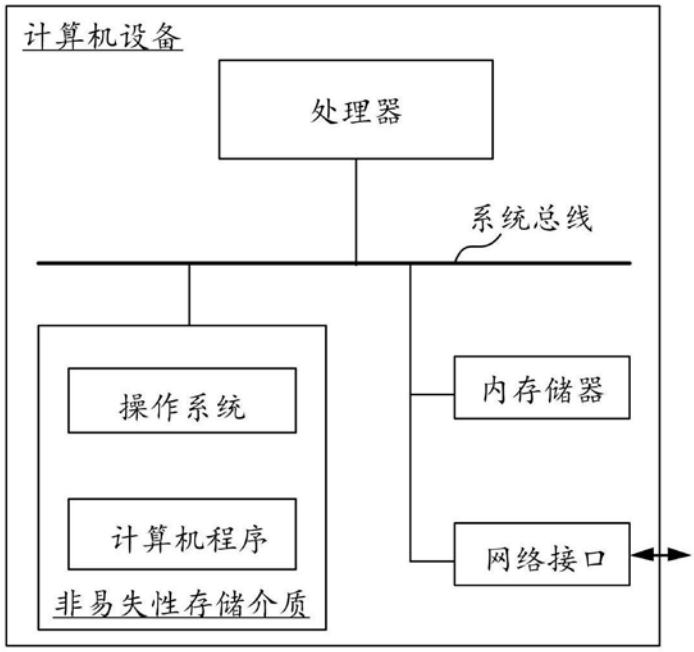


图11