



# (12)发明专利申请

(10)申请公布号 CN 106354548 A

(43)申请公布日 2017. 01. 25

(21)申请号 201610777613.5

(22)申请日 2016.08.31

(71)申请人 天津南大通用数据技术股份有限公司

地址 300384 天津市西青区华苑产业区海泰发展六道6号海泰绿色产业基地J-518

(72)发明人 武新 崔维力 张绍勇

(51)Int.Cl.

G06F 9/455(2006.01)

G06F 9/50(2006.01)

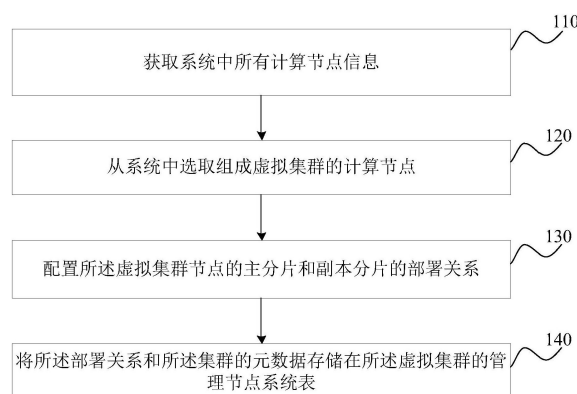
权利要求书2页 说明书5页 附图2页

## (54)发明名称

分布式数据库系统中虚拟集群创建、管理方法及装置

## (57)摘要

本发明提供了一种分布式数据库系统中虚拟集群创建、管理方法及装置,其中,所述方法包括:获取系统中所有计算节点信息;从系统中选取组成虚拟集群的计算节点;配置所述虚拟集群节点的主分片和副本分片的部署关系;将所述部署关系和所述集群的元数据存储在所述虚拟集群的管理节点系统表。可以根据业务系统将物理集群规划为虚拟集群,无需维护多套系统,有效降低部署和维护成本。



1. 一种分布式数据库系统中虚拟集群的创建方法,其特征在于,包括:  
获取系统中所有计算节点信息;  
从系统中选取组成虚拟集群的计算节点;  
配置所述虚拟集群节点的主分片和副本分片的部署关系;  
将所述部署关系和所述集群的元数据存储在该虚拟集群的管理节点系统表。
2. 一种分布式数据库系统中虚拟集群的管理方法,其特征在于,包括:  
在所述虚拟集群的管理节点创建表;  
在其它管理节点和虚拟集群的节点中创建表;  
管理集群节点的元数据系统表和虚拟集群的对应关系。
3. 根据权利要求2所述的方法,其特征在于:所述管理集群节点的元数据系统表和虚拟集群的对应关系,包括:  
从所述虚拟集群的管理节点读取元数据系统表;  
从元数据系统表中读取所述系统表对应的虚拟集群;  
从所述元数据系统表中读取节点上主分片和副分片数据的部署关系以及主分片和副分片的哈希值;  
计算加载数据的哈希值,并根据所述哈希值对应的数据主副分片所属的节点,将所述数据插入到所述节点的表中。
4. 根据权利要求2所述的方法,其特征在于,所述管理集群节点的元数据系统表和虚拟集群的对应关系,包括:  
通过虚拟集群的管理节点读取数据所属的虚拟集群;  
读取并返回所述虚拟集群表中的信息。
5. 根据权利要求4所述的方法,其特征在于,还包括:  
如果数据所属的虚拟集群不小于两个,则在第一虚拟集群上创建中间表,并将其它虚拟集群的查询结果重分布到所述中间表中。
6. 根据权利要求2所述的方法,其特征在于,所述管理集群节点的元数据系统表和虚拟集群的对应关系,包括:  
从所述虚拟集群的管理节点读取元数据系统表;  
锁住所述虚拟集群;  
根据修改的节点的数据确定虚拟集群节点上主分片和副分片数据的部署关系;  
对虚拟集群的数据重新分布。
7. 根据权利要求6所述的方法,其特征在于,所述方法还包括:将修改后的虚拟集群的元数据储存在管理集群的节点的系统表中。
8. 根据权利要求2所述的方法,其特征在于,包括:  
锁住移动节点对应的移出虚拟集群和移入虚拟集群;  
移出虚拟集群删除所述移动节点;  
移入虚拟集群增加所述移动节点;  
释放集群锁,移出虚拟集群和移入虚拟集群重新生成元数据系统表。
9. 根据权利要求2所述的方法,其特征在于,包括:  
判断镜像虚拟集群和复制虚拟集群结构是否相同;

在所述镜像虚拟集群和复制虚拟集群结构相同时,确定镜像虚拟集群和复制虚拟集群结构的区别数据;

将区别数据同步和复制至虚拟集群中。

10.一种分布式数据库系统中虚拟集群的创建装置,其特征在于,包括:

信息获取单元,用于获取系统中所有计算节点信息;

节点选取单元,用于从系统中选取组成虚拟集群的计算节点;

关系配置单元,用于配置所述虚拟集群节点的主分片和副本分片的部署关系;

存储单元,用于将所述部署关系和所述集群的元数据存储在该所述虚拟集群的管理节点系统表。

包括:表创建单元,用于在该所述虚拟集群的管理节点创建表;

创建单元,用于在其它管理节点和虚拟集群的节点中创建表;

关系管理单元,用于管理集群节点的元数据系统表和虚拟集群的对应关系。

## 分布式数据库系统中虚拟集群创建、管理方法及装置

### 技术领域

[0001] 本发明属于数据库技术领域,尤其是涉及一种分布式数据库系统中虚拟集群创建、管理方法及装置。

### 背景技术

[0002] 虚拟集群由多个客户虚拟机构成,这些客户虚拟机安装在由一个或多个物理集群构成的分布式服务器上。在逻辑上,处于一个虚拟集群的客户虚拟机通过一个跨越了多个物理网络的虚拟网络互连在一起。虚拟集群具有明确的边界。一个虚拟集群可以由多台物理机器构成,也可以由一个可运行在多个物理集群上的虚拟机主机操作系统管理物理机器的资源,虚拟机运行其上,并且可以运行与主机相异的操作系统。

[0003] 使用虚拟机的目的是合并同一台物理服务器的多个功能。这可以显著提高服务器的资源利用率与应用的灵活性。虚拟机可以在多个物理服务器上备份,以提高分布式并行度、容错性,加快灾难恢复速度。虚拟集群的节点数可以动态增减,与P2P网络中覆盖网络的规模变化类似。物理节点的失效会使得运行在其上的虚拟机也失效,但是虚拟机的失效不会影响主机系统。

[0004] 随着系统虚拟化的广泛应用,对运行在大量物理计算节点上的虚拟机(也称为虚拟集群)需要进行有效管理,从而构建高性能的虚拟计算环境。具体来讲,包括虚拟集群的部署、大规模集群的监视和管理,以及资源调度、负载均衡、服务器合并、容错等技术。

### 发明内容

[0005] 本发明提供了一种分布式数据库系统中虚拟集群的创建、管理方法及装置,以实现

对分布式数据库中虚拟集群的有效管理。

[0006] 第一方面,本发明实施例提供了一种分布式数据库系统中虚拟集群的创建方法,包括:

[0007] 获取系统中所有计算节点信息;

[0008] 从系统中选取组成虚拟集群的计算节点;

[0009] 配置所述虚拟集群节点的主分片和副本分片的部署关系;

[0010] 将所述部署关系和所述集群的元数据存储在该虚拟集群的管理节点系统表。

[0011] 第二方面,本发明实施例还提供了一种分布式数据库系统中虚拟集群的管理方法,包括:

[0012] 在所述虚拟集群的管理节点创建表;

[0013] 在其它管理节点和虚拟集群的节点中创建表;

[0014] 管理集群节点的元数据系统表和虚拟集群的对应关系。

[0015] 进一步的,所述管理集群节点的元数据系统表和虚拟集群的对应关系,包括:

[0016] 从所述虚拟集群的管理节点读取元数据系统表;

[0017] 从元数据系统表中读取所述系统表对应的虚拟集群;

- [0018] 从所述元数据系统表中读取节点上主分片和副分片数据的部署关系以及主分片和副分片的哈希值；
- [0019] 计算加载数据的哈希值,并根据所述哈希值对应的数据主副分片所属的节点,将所述数据插入到所述节点的表中。
- [0020] 进一步的,所述管理集群节点的元数据系统表和虚拟集群的对应关系,包括:
- [0021] 通过虚拟集群的管理节点读取数据所属的虚拟集群;
- [0022] 读取并返回所述虚拟集群表中的信息。
- [0023] 进一步的,还包括:
- [0024] 如果数据所属的虚拟集群不小于两个,则在第一虚拟集群上创建中间表,并将其它虚拟集群的查询结果重分布到所述中间表中。
- [0025] 进一步的,所述管理集群节点的元数据系统表和虚拟集群的对应关系,包括:
- [0026] 从所述虚拟集群的管理节点读取元数据系统表;
- [0027] 锁住所述虚拟集群;
- [0028] 根据修改的节点的数据确定虚拟集群节点上主分片和副分片数据的部署关系;
- [0029] 进一步的,所述方法还包括:将修改后的虚拟集群的元数据储存在管理集群的节点的系统表中。
- [0030] 对虚拟集群的数据重新分布。
- [0031] 第三方面,本发明实施例还提供了一种分布式数据库系统中虚拟集群的创建装置,包括:
- [0032] 信息获取单元,用于获取系统中所有计算节点信息;
- [0033] 节点选取单元,用于从系统中选取组成虚拟集群的计算节点;
- [0034] 关系配置单元,用于配置所述虚拟集群节点的主分片和副本分片的部署关系;
- [0035] 存储单元,用于将所述部署关系和所述集群的元数据存储在所述虚拟集群的管理节点系统表。
- [0036] 第四方面,本发明实施例还提供了一种分布式数据库系统中虚拟集群的管理方法,包括:
- [0037] 表创建单元,用于在所述虚拟集群的管理节点创建表;
- [0038] 创建单元,用于在其它管理节点和虚拟集群的节点中创建表;
- [0039] 关系管理单元,用于管理集群节点的元数据系统表和虚拟集群的对应关系。
- [0040] 本发明实施例提供的分布式数据库系统中虚拟集群创建、管理方法及装置,可以根据业务系统将物理集群规划为虚拟集群,无需维护多套系统,有效降低部署和维护成本。

## 附图说明

[0041] 为了更清楚地说明本发明实施例的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图

[0042] 图1是本发明实施例一提供的分布式数据库系统中虚拟集群的创建方法的流程示意图;

[0043] 图2是本发明实施例二提供的分布式数据库系统中虚拟集群的管理方法的流程图;

[0044] 图3是本发明实施例三提供的分布式数据库系统中虚拟集群的创建装置的结构示意图;

[0045] 图4是本发明实施例四提供的分布式数据库系统中虚拟集群的管理装置的结构示意图。

## 具体实施方式

[0046] 为了更好的理解本发明,下面结合具体实施例和附图对本发明进行进一步的描述。

[0047] 下面将结合本发明实施例中的附图,对本发明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发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,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,都属于本发明保护的范围。

### [0048] 实施例一

[0049] 图1为本发明实施例一提供的分布式数据库系统中虚拟集群的创建方法的流程图,本实施例可适用于分布式数据库系统中创建虚拟集群的情况,该方法可以由适用于分布式数据库系统中虚拟集群的创建装置来执行,该装置可由软件/硬件方式实现,并可集成于分布式数据库系统中。

[0050] 参见图1,所述分布式数据库系统中虚拟集群的创建方法,包括:

[0051] S110,获取系统中所有计算节点信息。

[0052] 遍历所述分布式数据系统,获取所有节点的信息。示例性的,通过虚拟集群的管理节点作为入口,获取所有节点的信息。

[0053] S120,从系统中选取组成虚拟集群的计算节点。

[0054] 从系统中的所有节点中选取虚拟集群所需要的计算节点。

[0055] S130,配置所述虚拟集群节点的主分片和副本分片的部署关系。

[0056] 设置数据主分片和副本分片对应的哈希值和选取的虚拟集群中计算节点的关系,配置数据主分片和副本分片的存储位置。

[0057] S140,将所述部署关系和所述集群的元数据存储在该虚拟集群的管理节点系统表。

[0058] 将步骤S130确定的虚拟集群节点的主分片和副本分片的部署关系和元数据存储在该虚拟集群的管理节点的系统表中,通过管理节点可获取虚拟集群的相应存储信息。

[0059] 本实施例可以根据业务系统将物理集群规划为虚拟集群,无需维护多套系统,有效降低部署和维护成本。

### [0060] 实施例二

[0061] 图2为本发明实施例二提供的分布式数据库系统中虚拟集群的管理方法的流程图,本实施例可适用于分布式数据库系统对虚拟集群进行管理的情况,该方法可以由适用于分布式数据库系统中虚拟集群的管理装置来执行,该装置可由软件/硬件方式实现,并可集成于分布式数据库系统中。

[0062] 参见图2,所述分布式数据库系统中虚拟集群的管理方法,包括:

[0063] S210,在所述虚拟集群的管理节点创建表。

[0064] 在确定的虚拟集群的管理接单创建表,并指定表对应的虚拟集群的名称。

[0065] S220,在其它管理节点和虚拟集群的节点中创建表。

[0066] 在所有的管理节点和该虚拟集群对应的计算节点上都创建表。

[0067] S230,管理集群节点的元数据系统表和虚拟集群的对应关系。

[0068] 将虚拟集群节点的主分片和副本分片的部署关系和元数据存储虚拟集群的管理节点的系统表的表中,并同步到虚拟集群的计算节点的表中,通过管理节点可获取虚拟集群的相应存储信息。

[0069] 示例性的,可以实现虚拟集群的数据加载:从所述虚拟集群的管理节点读取元数据系统表;从元数据系统表中读取所述系统表对应的虚拟集群;从所述元数据系统表中读取节点上主分片和副分片数据的部署关系以及主分片和副分片的哈希值;计算加载数据的哈希值,并根据所述哈希值对应的数据主副分片所属的节点,将所述数据插入到所述节点的表中。

[0070] 此外,还可实现对虚拟集群数据的读取,这里分为两种情况,读取的数据存储在一个或者多个虚拟集群中,在存储在一个集群中时,通过虚拟集群的管理节点读取数据所属的虚拟集群;读取并返回所述虚拟集群表中的信息。如果数据所属的虚拟集群不小于两个,则在第一虚拟集群上创建中间表,并将其它虚拟集群的查询结果重分布到所述中间表中。

[0071] 此外,也可实现增加和删除虚拟集群中计算节点,示例性的,从所述虚拟集群的管理节点读取元数据系统表;锁住所述虚拟集群;根据修改的节点的数据确定虚拟集群节点上主分片和副分片数据的部署关系;对虚拟集群的数据重新分布。在重分布完成后,释放集群锁,使得增加和删除计算节点后的虚拟集群可以重新提供服务。

[0072] 本实施例还可实现计算节点在不同的虚拟集群中互相移动,示例性的,包括:锁住移动节点对应的移出虚拟集群和移入虚拟集群;移出虚拟集群删除所述移动节点;移入虚拟集群增加所述移动节点;释放集群锁,移出虚拟集群和移入虚拟集群重新生成元数据系统表。在完成移动后,释放集群锁,使得移出虚拟集群和移入虚拟集群可以对外提供服务。

[0073] 本实施例还可实现对不同的虚拟集群互相复制,示例性的,包括:判断镜像虚拟集群和复制虚拟集群结构是否相同;所述镜像虚拟集群和复制虚拟集群结构相同时,确定镜像虚拟集群和复制虚拟集群结构的区别数据;将区别数据同步和复制至虚拟集群中。

[0074] 本实施例提供的分布式数据库系统中虚拟集群的管理方法,可以根据业务系统将物理集群规划为虚拟集群,并对虚拟集群的各种业务进行有效管理,无需维护多套系统,有效降低部署和维护成本。

[0075] 实施例三

[0076] 图3是本发明实施例三提供的分布式数据库系统中虚拟集群的创建装置的结构示意图,如图3所示,所述装置包括:

[0077] 信息获取单元310,用于获取系统中所有计算节点信息;

[0078] 节点选取单元320,用于从系统中选取组成虚拟集群的计算节点;

[0079] 关系配置单元330,用于配置所述虚拟集群节点的主分片和副本分片的部署关系;

[0080] 存储单元340,用于将所述部署关系和所述集群的元数据存储所述虚拟集群的

管理节点系统表。

[0081] 本实施例提供的分布式数据库系统中虚拟集群创建装置,可以根据业务系统将物理集群规划为虚拟集群,无需维护多套系统,有效降低部署和维护成本。

[0082] 实施例四

[0083] 图4是本发明实施例四提供的分布式数据库系统中虚拟集群的管理装置的结构示意图,如图4所示,所述装置包括:

[0084] 表创建单元410,用于在所述虚拟集群的管理节点创建表;

[0085] 创建单元420,用于在其它管理节点和虚拟集群的节点中创建表;

[0086] 关系管理单元430,用于管理集群节点的元数据系统表和虚拟集群的对应关系。

[0087] 本实施例提供的分布式数据库系统中虚拟集群的管理装置,可以根据业务系统将物理集群规划为虚拟集群,并对虚拟集群的各种业务进行有效管理,无需维护多套系统,有效降低部署和维护成本。

[0088] 本领域普通技术人员可以理解:实现上述各方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成。前述的程序可以存储于一计算机可读取存储介质中。该程序在执行时,执行包括上述各方法实施例的步骤;而前述的存储介质包括:ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0089] 最后应说明的是:以上各实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述各实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分或者全部技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

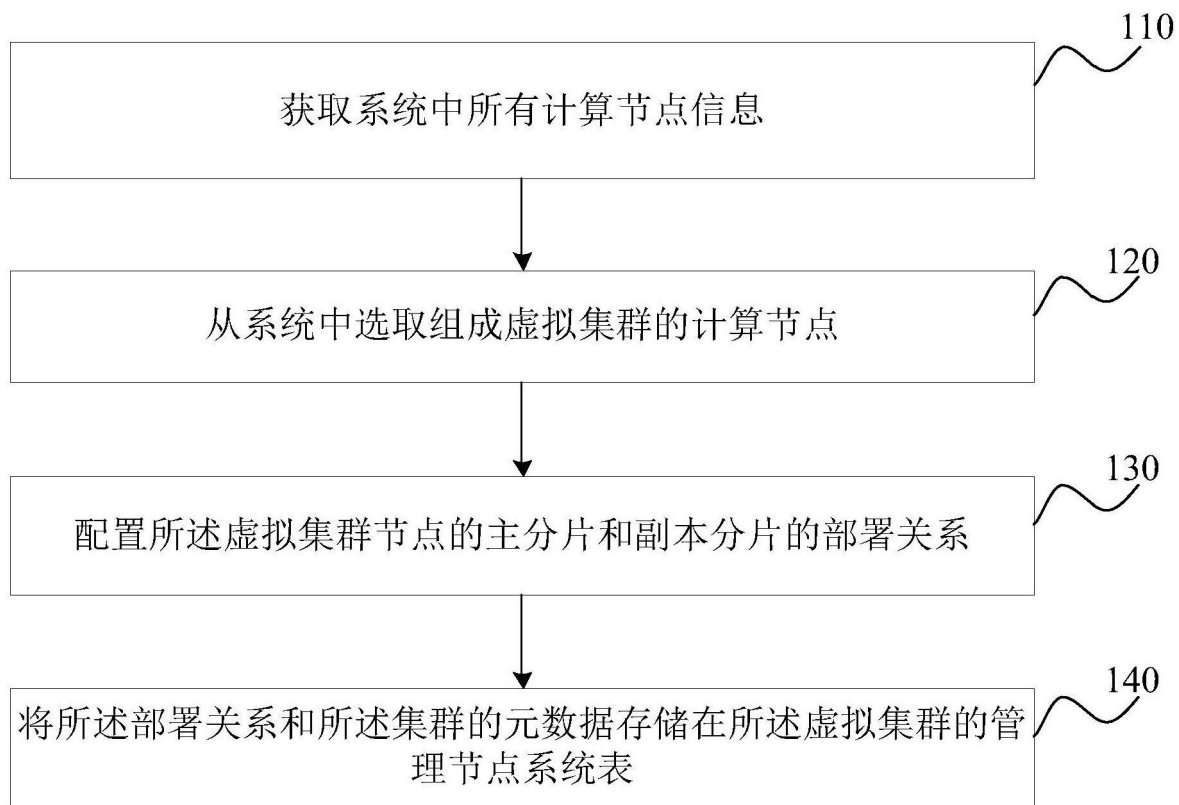


图1

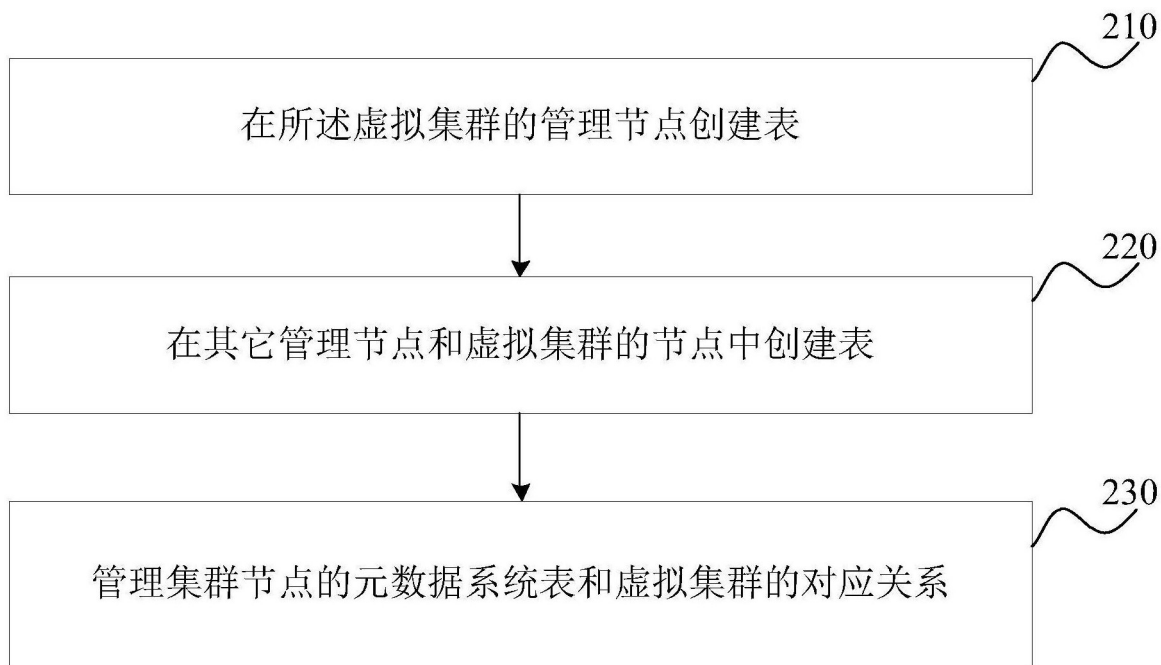


图2

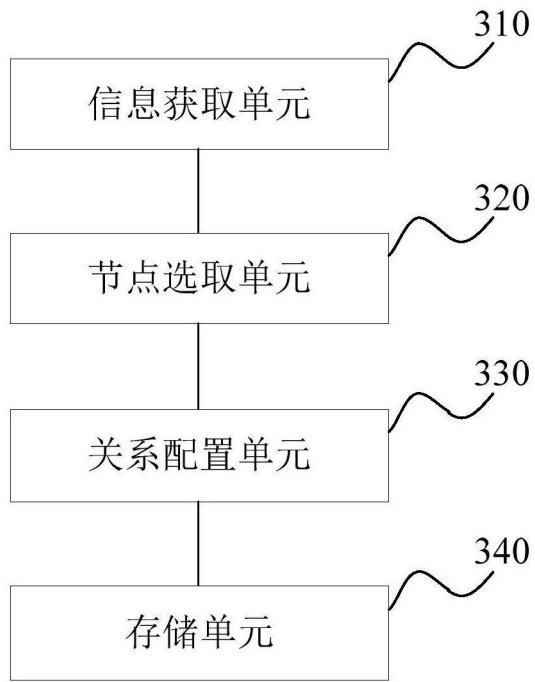


图3

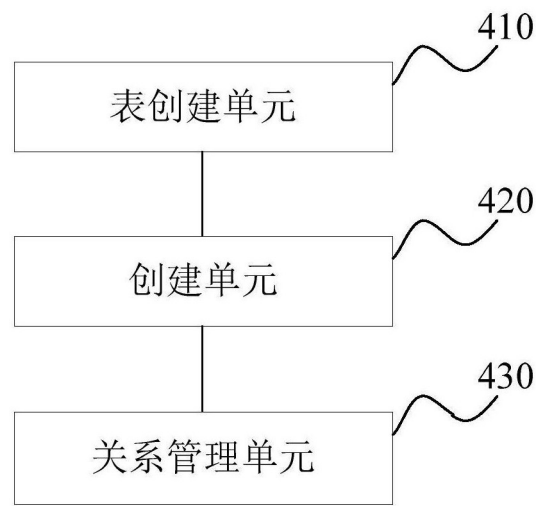


图4