



(12) 发明专利申请

(10) 申请公布号 CN 120017618 A

(43) 申请公布日 2025. 05. 16

(21) 申请号 202510173932.4

(22) 申请日 2025.02.17

(71) 申请人 北京火山引擎科技有限公司

地址 100190 北京市海淀区紫金数码园4号楼13层1309

(72) 发明人 任峻驰 蔡乐 汪阳 王剑

(74) 专利代理机构 北京三聚阳光知识产权代理有限公司 11250

专利代理师 刘贺秋

(51) Int. Cl.

H04L 49/25 (2022.01)

H04L 49/111 (2022.01)

H04L 49/60 (2022.01)

H04L 49/65 (2022.01)

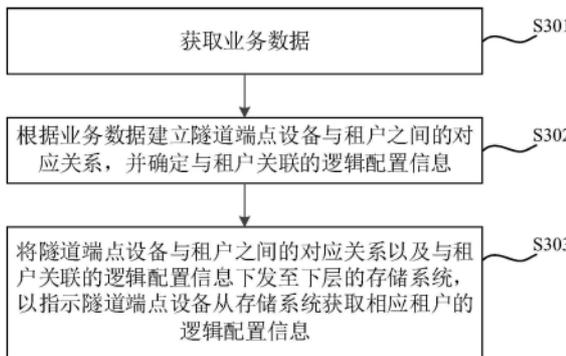
权利要求书2页 说明书12页 附图5页

(54) 发明名称

软件定义网络的配置方法、装置、设备、介质及网络

(57) 摘要

本公开涉及软件定义网络技术领域,公开了软件定义网络的配置方法、装置、设备、介质及网络,该方法包括:获取业务数据;业务数据包括各个租户的虚拟资源对应的逻辑配置信息以及虚拟资源对应的隧道端点设备;根据业务数据建立隧道端点设备与租户之间的对应关系,并确定与租户关联的逻辑配置信息;将隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息下发至下层的存储系统,以指示隧道端点设备从存储系统获取相应租户的逻辑配置信息。本公开使得各个隧道端点设备的第二控制器可以从存储系统中只获取自身所需的配置,不需要向每台隧道端点设备下发全量租户的配置,能够有效降低各个隧道端点设备的配置数量。



1. 一种软件定义网络的配置方法,其特征在于,应用于软件定义网络的第一控制器,所述方法包括:

获取业务数据;所述业务数据包括各个租户的虚拟资源对应的逻辑配置信息以及所述虚拟资源对应的隧道端点设备;

根据所述业务数据建立隧道端点设备与租户之间的对应关系,并确定与租户关联的逻辑配置信息;

将所述隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息下发至下层的存储系统,以指示隧道端点设备从所述存储系统获取相应租户的逻辑配置信息。

2. 根据权利要求1所述的方法,其特征在于,所述隧道端点设备与租户之间的对应关系所对应的索引结构,包括:第一层的隧道端点设备的设备标识,以及第二层的租户标识。

3. 根据权利要求1或2所述的方法,其特征在于,所述确定与租户关联的逻辑配置信息,包括:

确定键值存储形式的与租户关联的逻辑配置信息;所述与租户关联的逻辑配置信息的键包括租户标识。

4. 一种软件定义网络的配置方法,其特征在于,应用于隧道端点设备的第二控制器,所述方法包括:

根据上层的存储系统中存储的隧道端点设备与租户之间的对应关系,查询与本地隧道端点设备相对应的目标租户;

从所述存储系统获取与所述目标租户相关联的目标逻辑配置信息;所述存储系统存储有与租户关联的逻辑配置信息;

根据所述目标逻辑配置信息生成相应的流表表项,并下发所述流表表项。

5. 根据权利要求4所述的方法,其特征在于,所述隧道端点设备与租户之间的对应关系所对应的索引结构,包括:第一层的隧道端点设备的设备标识,以及第二层的租户标识;

所述根据上层的存储系统中存储的隧道端点设备与租户之间的对应关系,查询与本地隧道端点设备相对应的目标租户,包括:

根据本地隧道端点设备的设备标识对存储系统中存储的隧道端点设备与租户之间的对应关系所对应的索引结构进行前缀匹配查询,确定目标租户的租户标识。

6. 根据权利要求4或5所述的方法,其特征在于,所述存储系统包括存储和通知中间件;所述与租户关联的逻辑配置信息采用键值存储形式,且所述与租户关联的逻辑配置信息的键包括租户标识;

所述从所述存储系统获取与所述目标租户相关联的目标逻辑配置信息,包括:

向所述存储和通知中间件发起监听请求;所述监听请求用于监听所述目标租户的租户标识;

获取在所述目标租户的租户标识所对应的值发生变化时,所述存储和通知中间件推送的监听事件;所述监听事件包括与所述目标租户相关联的目标逻辑配置信息。

7. 一种软件定义网络,其特征在于,包括:第一控制器、存储系统和多个隧道端点设备;所述隧道端点设备设有第二控制器;

所述第一控制器用于执行权利要求1至3中任一项所述的软件定义网络的配置方法;

所述第二控制器用于执行权利要求4至6中任一项所述的软件定义网络的配置方法。

8. 一种软件定义网络的配置装置,其特征在於,应用于软件定义网络的第一控制器,所述装置包括:

数据获取模块,用于获取业务数据;所述业务数据包括各个租户的虚拟资源对应的逻辑配置信息以及所述虚拟资源对应的隧道端点设备;

处理模块,用于根据所述业务数据建立隧道端点设备与租户之间的对应关系,并确定与租户关联的逻辑配置信息;

配置下发模块,用于将所述隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息下发至下层的存储系统,以指示隧道端点设备从所述存储系统获取相应租户的逻辑配置信息。

9. 一种软件定义网络的配置装置,其特征在於,应用于隧道端点设备的第二控制器,所述装置包括:

查询模块,用于根据上层的存储系统中存储的隧道端点设备与租户之间的对应关系,查询与本地隧道端点设备相对应的目标租户;

配置获取模块,用于从所述存储系统获取与所述目标租户相关联的目标逻辑配置信息;所述存储系统存储有与租户关联的逻辑配置信息;

流表下发模块,用于根据所述目标逻辑配置信息生成相应的流表表项,并下发所述流表表项。

10. 一种计算机设备,其特征在於,包括:

存储器和处理器,所述存储器和所述处理器之间互相通信连接,所述存储器中存储有计算机指令,所述处理器通过执行所述计算机指令,从而执行权利要求1至6中任一项所述的软件定义网络的配置方法。

11. 一种计算机可读存储介质,其特征在於,所述计算机可读存储介质上存储有计算机指令,所述计算机指令用于使计算机执行权利要求1至6中任一项所述的软件定义网络的配置方法。

## 软件定义网络的配置方法、装置、设备、介质及网络

### 技术领域

[0001] 本公开涉及软件定义网络技术领域,具体涉及软件定义网络的配置方法、装置、设备、介质及网络。

### 背景技术

[0002] 随着云计算和虚拟化技术的发展,虚拟机和容器相比于物理机提供了更高的灵活性和资源利用率。Overlay网络(覆盖网络)通过VXLAN(Virtual eXtensible Local Area Network,虚拟可扩展局域网)技术在物理网络之上创建虚拟网络,允许跨主机的虚拟机和容器高效通信。当前,SDN(Software-defined Networking,软件定义网络)技术是一种常用的网络虚拟化的实现方式,搭配虚拟交换机,实现网络虚拟化的能力。

[0003] 在SDN架构下,传统的SDN控制器会为每台VTEP设备下发了全量租户的流表,但随着租户数量逐渐增多,会导致各个VTEP设备的流表表项增多,导致SDN控制器和VTEP设备处理压力增大。

### 发明内容

[0004] 有鉴于此,本公开提供了一种软件定义网络的配置方法、装置、设备、介质及网络,以解决控制器处理压力较大的问题。

[0005] 第一方面,本公开提供了一种软件定义网络的配置方法,应用于软件定义网络的第一控制器,所述方法包括:

[0006] 获取业务数据;所述业务数据包括各个租户的虚拟资源对应的逻辑配置信息以及所述虚拟资源对应的隧道端点设备;

[0007] 根据所述业务数据建立隧道端点设备与租户之间的对应关系,并确定与租户关联的逻辑配置信息;

[0008] 将所述隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息下发至下层的存储系统,以指示隧道端点设备从所述存储系统获取相应租户的逻辑配置信息。

[0009] 第二方面,本公开提供了一种软件定义网络的配置方法,应用于隧道端点设备的第二控制器,所述方法包括:

[0010] 根据上层的存储系统中存储的隧道端点设备与租户之间的对应关系,查询与本地隧道端点设备相对应的目标租户;

[0011] 从所述存储系统获取与所述目标租户相关联的目标逻辑配置信息;所述存储系统存储有与租户关联的逻辑配置信息;

[0012] 根据所述目标逻辑配置信息生成相应的流表表项,并下发所述流表表项。

[0013] 第三方面,本公开提供了一种软件定义网络,包括:第一控制器、存储系统和多个隧道端点设备;所述隧道端点设备设有第二控制器;

[0014] 所述第一控制器用于执行上述第一方面或其对应的任一实施方式的软件定义网

络的配置方法；

[0015] 所述第二控制器用于执行上述第二方面或其对应的任一实施方式的软件定义网络的配置方法。

[0016] 第四方面,本公开提供了一种软件定义网络的配置装置,应用于软件定义网络的第一控制器,所述装置包括:

[0017] 数据获取模块,用于获取业务数据;所述业务数据包括各个租户的虚拟资源对应的逻辑配置信息以及所述虚拟资源对应的隧道端点设备;

[0018] 处理模块,用于根据所述业务数据建立隧道端点设备与租户之间的对应关系,并确定与租户关联的逻辑配置信息;

[0019] 配置下发模块,用于将所述隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息下发至下层的存储系统,以指示隧道端点设备从所述存储系统获取相应租户的逻辑配置信息。

[0020] 第五方面,本公开提供了一种软件定义网络的配置装置,应用于隧道端点设备的第二控制器,所述装置包括:

[0021] 查询模块,用于根据上层的存储系统中存储的隧道端点设备与租户之间的对应关系,查询与本地隧道端点设备相对应的目标租户;

[0022] 配置获取模块,用于从所述存储系统获取与所述目标租户相关联的目标逻辑配置信息;所述存储系统存储有与租户关联的逻辑配置信息;

[0023] 流表下发模块,用于根据所述目标逻辑配置信息生成相应的流表表项,并下发所述流表表项。

[0024] 第六方面,本公开提供了一种计算机设备,包括:存储器和处理器,存储器和处理器之间互相通信连接,存储器中存储有计算机指令,处理器通过执行计算机指令,从而执行上述第一方面或其对应的任一实施方式的软件定义网络的配置方法。

[0025] 第七方面,本公开提供了一种计算机可读存储介质,该计算机可读存储介质上存储有计算机指令,计算机指令用于使计算机执行上述第一方面或其对应的任一实施方式的软件定义网络的配置方法。

[0026] 第八方面,本公开提供了一种计算机程序产品,包括计算机指令,计算机指令用于使计算机执行上述第一方面或其对应的任一实施方式的软件定义网络的配置方法。

[0027] 本公开中的第一控制器将上层的业务数据转换为隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息,并将这些信息下发至存储系统,使得各个隧道端点设备的第二控制器可以从存储系统中只获取自身所需的配置,不需要向每台隧道端点设备下发全量租户的配置,能够有效降低各个隧道端点设备的配置数量。并且,利用存储系统实现两层控制器的逻辑解耦,能够提高整体的稳定性和可靠性。

[0028] 并且,本地隧道端点设备的第二控制器对存储系统中的隧道端点设备与租户之间的对应关系进行查询,可以确定与自身对应的目标租户,进而从存储系统中获取到与目标租户相关联的目标逻辑配置信息,能够实现按需下发网络配置,本地隧道端点设备只需要获取必须的目标逻辑配置信息,不需要获取全量租户的配置,能够减少每个隧道端点设备的配置数量,降低每个隧道端点设备的负载,减轻处理压力。

## 附图说明

[0029] 为了更清楚地说明本公开具体实施方式或相关技术中的技术方案,下面将对具体实施方式或相关技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本公开的一些实施方式,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。

[0030] 图1是传统SDN的一种架构示意图;

[0031] 图2是根据本公开实施例的软件定义网络的一种架构示意图;

[0032] 图3是根据本公开实施例的软件定义网络的配置方法的流程示意图;

[0033] 图4是根据本公开实施例的第一控制器的一种逻辑架构示意图;

[0034] 图5是根据本公开实施例的另一软件定义网络的配置方法的流程示意图;

[0035] 图6是根据本公开实施例的第二控制器的一种逻辑架构示意图;

[0036] 图7是根据本公开实施例的第一控制器的配置装置的结构框图;

[0037] 图8是根据本公开实施例的第二控制器的配置装置的结构框图;

[0038] 图9是本公开实施例的计算机设备的硬件结构示意图。

## 具体实施方式

[0039] 为使本公开实施例的目的、技术方案和优点更加清楚,下面将结合本公开实施例中的附图,对本公开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公开一部分实施例,而不是全部的实施例。基于本公开中的实施例,本领域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,都属于本公开保护的范围。

[0040] SDN(Software-Defined Networking,软件定义网络)是一种新兴的网络架构模式,是一种网络虚拟化和容器化方法。传统网络设备如交换机和路由器,其数据平面(负责实际转发数据包)和控制平面(决定数据如何流动)在硬件中紧密集成,这种耦合状态使得网络配置和管理变得复杂且缺乏灵活性。而SDN的核心理念在于将控制平面从网络硬件中分离出来,交由SDN控制器进行统一管理,这样能够优化网络资源并快速调整网络,使其适应不断变化的业务需求、应用和流量,从而实现更高效的网络控制与管理。

[0041] 图1示出了SDN的一种架构示意图。如图1所示,控制平面的SDN控制器是SDN网络的大脑,主要负责下发路由决策和策略,对整个网络流量进行管理。

[0042] 北向接口是控制器与上层应用平面的通信接口,允许应用程序通过API(应用程序接口)与网络设备进行交互,这使得网络管理和应用开发更加便捷,为网络的智能化和自动化提供了可能。

[0043] 南向接口为控制器与网络设备之间的通信接口,通过南向接口,控制器能够向网络设备发送指令,告知它们如何处理特定的数据流,确保数据按照预定的规则进行传输。南向接口采用的通信协议一般是OpenFlow(一种网络通信协议),它允许SDN控制器直接与网络设备(如交换机、路由器)进行通信,从而实现对网络流量的集中管理和动态配置。通过OpenFlow,SDN控制器可以下发指令,控制网络设备如何处理和转发数据包。

[0044] 数据平面可以包括传统的网络设备(如交换机、路由器),其主要职责是负责数据包的转发工作。在SDN架构下,数据平面只需按照SDN控制器的指令进行简单的数据转发操作,无需再进行复杂的路由决策。

[0045] 随着云计算和虚拟化技术的不断发展,网络虚拟化成为了构建灵活、可扩展网络架构的关键技术之一。Open vSwitch (OVS) 作为一种功能强大的开源虚拟交换机,被广泛应用于云计算和虚拟化环境中。

[0046] OVS是一个应用于软件定义网络中的支持多层数据转发的高质量虚拟交换机,其基于软件定义网络 (SDN) 的思想,通过软件实现传统硬件交换机的功能。通过虚拟化技术,为每个虚拟机提供一个独立的网络接口,从而实现在同一物理主机上的多个虚拟机之间的网络隔离和通信。

[0047] VXLAN是一种网络虚拟化技术,VTEP设备是VXLAN网络的边缘设备,是VXLAN隧道的起点和终点,通常是物理或虚拟交换机、服务器或其他支持VXLAN的网络设备。VTEP设备的工作是在彼此之间创建和终止隧道,负责VXLAN协议报文的封包和解包,也就是在虚拟报文上封装VTEP通信的报文头部,能够实现不同VTEP设备中虚拟机之间的通信。

[0048] 目前,在SDN架构下,所有网络配置都是静态下发的,例如流表、VTEP (VXLAN Tunnel Endpoints, VXLAN隧道端点) 设备信息等。传统的SDN控制器会为每台VTEP设备下发了全量租户的流表。然而随着网络内的VPC (Virtual Private Cloud, 虚拟私有云) 租户 (Tenant) 逐渐增多,不同租户的虚拟机和容器也逐渐变多,集群规模逐渐增大,从而导致各个设备的流表表项会逐渐增多,导致SDN控制器和虚拟交换机处理压力增大,还会导致热升级、故障等场景恢复时间变长等问题。

[0049] 具体而言,传统的SDN架构分布式部署,各个VTEP设备资源管理加锁机制是必要的,以防止数据不一致或冲突;但大规模集群下加锁竞争会变得更加激烈,导致性能恶化。由于每台VTEP设备下发全量租户的配置,会造成每台VTEP设备可能存在大量无效的配置,因为并非所有租户都会使用到所有VTEP设备,这些无效配置不仅占用了设备资源,还降低了整个集群的性能。此外,若采用单机单点部署模式,如果控制器发生故障或受到攻击,整个SDN网络可能会受到影响,控制容灾能力较低,难以满足高可用性要求。

[0050] 根据本公开实施例,提供了一种软件定义网络的配置方法实施例,需要说明的是,在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行,并且,虽然在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤。

[0051] 在本实施例中提供了一种软件定义网络的配置方法,应用于软件定义网络中的控制器;其中,该软件定义网络内的控制器采用两层部署;并且,两层控制器之间设有存储系统,实现两层控制器的逻辑解耦,两层控制器不再直接交互,而是通过存储系统进行信息的存储和获取,两层控制器之间的耦合度大大降低,存储系统作为中间的缓冲层,可以有效解决集群内的事件风暴问题,提高了系统的稳定性和可维护性。

[0052] 图2示出了该软件定义网络的一种架构示意图,如图2所示,该软件定义网络包括:第一控制器、存储系统和多个隧道端点设备;所述隧道端点设备设有第二控制器。

[0053] 其中,第一控制器为集群内做单点部署的组件,其为上层的SDN控制器。该第一控制器与存储系统相连,用于向存储系统下发相应的配置信息。第二控制器为各个隧道端点设备中的控制代理组件 (Agent),其与存储系统相连,用于获取自身所需的配置信息。如图2所示,该隧道端点设备具体为VTEP设备。

[0054] 本实施例中提供的一种软件定义网络的配置方法,可应用于上述的第一控制器,

图3是根据本公开实施例的软件定义网络的配置方法的流程图,如图3所示,该流程包括以下步骤。

[0055] 步骤S301,获取业务数据;该业务数据包括各个租户的虚拟资源对应的逻辑配置信息以及虚拟资源对应的隧道端点设备。

[0056] 本实施例中,租户可以在应用平面创建自身所需的虚拟资源,该虚拟资源例如可以是虚拟机、容器等;并且,对该虚拟资源进行逻辑配置,可以生成相应的逻辑配置信息,例如,该逻辑配置信息可以包括:限速配置、弹性网卡配置、带宽配置等。

[0057] 并且,租户创建的虚拟资源会被调度至相应的节点,该节点所对应的设备为隧道端点设备,例如VTEP设备。在应用平面可以记录有各个租户的虚拟资源对应的逻辑配置信息以及虚拟资源对应的隧道端点设备。

[0058] 第一控制器可以从上层应用平面获取各个租户的业务数据,该业务数据即包括应用平面记录的租户的虚拟资源对应的逻辑配置信息以及虚拟资源对应的隧道端点设备。其中,第一控制器设有北向接口,可以基于该北向接口获取业务数据。

[0059] 步骤S302,根据业务数据建立隧道端点设备与租户之间的对应关系,并确定与租户关联的逻辑配置信息。

[0060] 本实施例中,为便于下层的各个隧道端点设备可以获取到自身所需的有效信息,第一控制器对业务数据进行了转换。

[0061] 具体地,根据该业务数据,可以确定各个隧道端点设备分别对应哪个或哪些租户(即隧道端点设备可以对应一个或多个租户),从而可以建立隧道端点设备与租户之间的对应关系。例如,租户A创建了一台虚拟机1,这台虚拟机1被调度到了节点B,该节点B对应VTEP设备2,则第一控制器可以生成VTEP设备2与租户A之间的对应关系。

[0062] 并且,根据该业务数据,可以为逻辑配置信息关联相应的租户,使得各个逻辑配置信息与租户是关联的。

[0063] 步骤S303,将隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息下发至下层的存储系统,以指示隧道端点设备从存储系统获取相应租户的逻辑配置信息。

[0064] 本实施例中,第一控制器所生成的信息并不直接下发至各个隧道端点设备,而是下发至下层的存储系统,由该存储系统记录隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息,即存储系统存储有全量租户的配置,后续各个隧道端点设备从该存储系统中获取相应的逻辑配置信息;其中,由于存储系统中存储的信息包括隧道端点设备与租户之间的对应关系,且各个逻辑配置信息与租户是关联的,使得各个隧道端点设备可以只获取本身所需的配置信息,不需要全量信息,从而能够减少各个隧道端点设备的配置数量,后续对隧道端点设备获取配置的过程进行说明。

[0065] 图4示出了第一控制器的一种逻辑架构示意图,该第一控制器作为集群内做单点部署的组件,其主要功能如下:

[0066] (1) 北向对外接口层:基于北向对外提供HTTP接口服务,可以向用户平面查询控制器的业务数据,例如虚拟机规格、设备带宽等数据。

[0067] (2) 北向对象模型和南向数据配置映射:北向对象模型为虚拟机模型等,通过对模型进行定义、对象模型属性校验,得到相应的北向对象数据,进而将北向对象数据转换为南

向数据配置,该南向数据配置为需要存储至存储系统的数据,例如隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息等。

[0068] (3) 南向逻辑处理层:负责南向数据批量下发,以发送至存储系统;并且,还用于处理南向设备上报信息,其中,该南向设备为隧道端点设备中的第二控制器,其上报的信息例如包括虚拟机状态等。

[0069] (4) 南向接口层:负责对接南向下发通道组件,例如存储系统的通道组件。

[0070] 本实施例提供的软件定义网络的配置方法,第一控制器将上层的业务数据转换为隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息,且将这些信息下发至存储系统,使得各个隧道端点设备的第二控制器可以从存储系统中只获取自身所需的配置,不需要向每台隧道端点设备下发全量租户的配置,能够有效降低各个隧道端点设备的配置数量。并且,利用存储系统实现两层控制器的逻辑解耦,能够提高整体的稳定性和可靠性。

[0071] 在一些可选的实施方式中,该隧道端点设备与租户之间的对应关系所对应的索引结构,包括:第一层的隧道端点设备的设备标识,以及第二层的租户标识。

[0072] 本实施例中,第一控制器生成的隧道端点设备与租户之间的对应关系,其按照一定的索引结构存储在该存储系统中。其中,该索引结构的第一层为隧道端点设备的设备标识,例如隧道端点设备的IP地址等,索引结构的第二层为租户标识,例如租户ID等。

[0073] 基于该索引结构存储隧道端点设备与租户之间的对应关系,有利于隧道端点设备基于各自的设备标识进行匹配查询,例如前缀匹配查询,从而快速获取到自身对应的租户标识,进而获取到与该租户标识相关联的逻辑配置信息,后续对该过程进行详细说明。

[0074] 可选地,上述步骤S302中“确定与租户关联的逻辑配置信息”的过程,可以包括:确定键值存储形式的与租户关联的逻辑配置信息;与租户关联的逻辑配置信息的键包括租户标识。

[0075] 本实施例中,该存储系统为键值存储系统,例如ETCD;其中,ETCD是一个开源的分布式键值存储系统,主要用于存储和管理配置信息、服务发现和协调分布式系统中的数据。ETCD中的ETC和D分别来源于unix“/etc”文件夹和分布式系统“Distributed”,“/etc”文件夹为单个系统存储配置数据的地方,ETCD存储大规模分布式系统的配置信息。

[0076] 第一控制器确定的与租户关联的逻辑配置信息,以键值存储形式对其进行存储,即存储系统中存储的租户关联的逻辑配置信息,是键值对(key-value)。其中,与租户关联的逻辑配置信息的键(key)至少包括租户标识,其值(value)为相应的逻辑配置信息。以键值存储形式存储与租户关联的逻辑配置信息,便于隧道端点设备基于租户标识对其进行监听,从而及时获取到自身所需且更新后的逻辑配置信息,后续对隧道端点设备的监听过程进行说明。

[0077] 本实施例中提供的另一种软件定义网络的配置方法,可应用于隧道端点设备中的控制器,即上述的第二控制器,图5是根据本公开实施例的软件定义网络的配置方法的流程图,如图5所示,该流程包括以下步骤。

[0078] 步骤S501,根据上层的存储系统中存储的隧道端点设备与租户之间的对应关系,查询与本地隧道端点设备相对应的目标租户。

[0079] 本实施例中,如上所述,存储系统中存储由第一控制器下发的全量租户的配置,具

体包括:隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息。

[0080] 其中,该隧道端点设备与租户之间的对应关系表示隧道端点设备与哪个或哪些租户相对应;对于执行该方法的本地隧道端点设备,可以根据该隧道端点设备与租户之间的对应关系进行查询,从而确定与本地隧道端点设备相对应的租户,为便于描述,将与本地隧道端点设备相对应的租户称为目标租户。

[0081] 可以理解,一般情况下,隧道端点设备只与部分租户之间具有对应关系,即只有这一部分租户会被作为目标租户;换句话说,与本地隧道端点设备无关的租户是无效租户,对于本地隧道端点设备,后续不对这些无效租户进行处理。

[0082] 步骤S502,从存储系统获取与目标租户相关联的目标逻辑配置信息;该存储系统存储有与租户关联的逻辑配置信息。

[0083] 本实施例中,第二控制器确定相对应的目标租户后,可以从存储系统中进一步获取与目标租户相关联的逻辑配置信息,即目标逻辑配置信息,该目标逻辑配置信息即为本地隧道端点设备所需的配置,使得下发到本地隧道端点设备的所有配置均是必要的,不包含无效租户的逻辑配置信息;即,本地隧道端点设备可以不保存全量租户的配置。

[0084] 步骤S503,根据目标逻辑配置信息生成相应的流表表项,并下发流表表项。

[0085] 本实施例中,第二控制器获取到必要的目标逻辑配置信息后,即可根据该目标逻辑配置信息计算得到相应的转发规则,从而生成流表表项(Flow Entry),每个流表表项可以定义一组匹配条件和相应的转发规则,最终下发该流表表项。如图2所示,第二控制器可以通过OpenFlow协议将流表表项下发至虚拟交换机,该虚拟交换机例如为OVS-dpdk(Data Plane Development Kit,一种数据平面开发套件),以基于该流表表项对相应的流量进行处理,实现数据转发。

[0086] 本实施例提供的软件定义网络的配置方法,本地隧道端点设备的第二控制器对存储系统中的隧道端点设备与租户之间的对应关系进行查询,可以确定与自身对应的目标租户,进而从存储系统中获取到与目标租户相关联的目标逻辑配置信息,能够实现按需下发网络配置,本地隧道端点设备只需要获取必须的目标逻辑配置信息,不需要获取全量租户的配置,能够减少每个隧道端点设备的配置数量,降低每个隧道端点设备的负载,减轻处理压力。

[0087] 在一些可选的实施方式中,如上所述,隧道端点设备与租户之间的对应关系所对应的索引结构,包括:第一层的隧道端点设备的设备标识,以及第二层的租户标识。

[0088] 并且,上述步骤S501“根据上层的存储系统中存储的隧道端点设备与租户之间的对应关系,查询与本地隧道端点设备相对应的目标租户”,包括步骤a1。

[0089] 步骤a1:根据本地隧道端点设备的设备标识对存储系统中存储的隧道端点设备与租户之间的对应关系进行前缀匹配查询,确定目标租户的租户标识。

[0090] 本实施例中,该对应关系的索引结果包括两层,分别为隧道端点设备的设备标识、租户标识。其中,该索引结构具体可以为存储路径,该存储路径为分层组织的目录,其第一层为隧道端点设备的设备标识,第二层为租户标识。具体地,该索引结构的格式可表示为: / <隧道端点设备的设备标识> / <租户标识>。

[0091] 例如,存储系统为ETCD等键值存储系统,该索引结构为对应关系的键,其值即可包括与相应隧道端点设备相对应的租户标识。或者,该索引结构第二层中的租户标识与具体

的数值相同,则基于该索引结构即可直接确定相对应的租户标识。

[0092] 具体地,本地隧道端点设备知道自身的设备标识,基于自身的设备标识可以对该索引结构进行前缀匹配查询,即对索引结构第一层的设备标识与自身的设备标识进行比对,从而能够查询到与自身设备标识相对应的各个租户标识,这些租户标识即为目标租户的租户标识。

[0093] 本实施例中,基于该索引结构存储隧道端点设备与租户之间的对应关系,本地隧道端点设备可以基于自身的设备标识对该索引结果进行前缀匹配查询,从而能够简单快速地确定自身设备下的各个目标租户,进而基于后续步骤获取到相关的目标逻辑配置信息。

[0094] 可选地,该存储系统包括存储和通知中间件;顾名思义,该存储和通知中间件除了具有存储功能之外,还具有通知功能;例如,该存储系统可以是ETCD,通过ETCD充当该存储和通知中间件。

[0095] 并且,该存储系统是键值存储系统,其中与租户关联的逻辑配置信息采用键值存储形式,且与租户关联的逻辑配置信息的键包括租户标识。其中,逻辑配置信息具体可以包括逻辑配置类型以及各个配置项对应的标识(例如ID),该与租户关联的逻辑配置信息的键具体可以包括:租户标识、逻辑配置类型以及配置项标识,其在存储系统中的存储格式例如为: /<逻辑配置类型>/<租户标识>/<配置项标识>。

[0096] 上述步骤S502“从存储系统获取与目标租户相关联的目标逻辑配置信息”,包括步骤b1至步骤b2。

[0097] 步骤b1,向存储和通知中间件发起监听请求;该监听请求用于监听目标租户的租户标识。

[0098] 步骤b2,获取在目标租户的租户标识所对应的值发生变化时,存储和通知中间件推送的监听事件;监听事件包括与目标租户相关联的目标逻辑配置信息。

[0099] 本实施例中,第二控制器确定目标租户的租户标识后,可以向该存储和通知中间件发起监听请求,以监听该目标租户的租户标识对应的值(value)的变化。在该租户标识对应的值发生创建、更新或删除等操作时,存储和通知中间件即可生成包括变化后的值的监听事件,变化后的值即包括目标租户关联的目标逻辑配置信息,之后向该第二控制器推送该监听事件,使得第二控制器可以基于监听的方式及时获取到所需的目标逻辑配置信息。

[0100] 图6示出了第二控制器的一种逻辑架构示意图,每个隧道端点设备均配置有第二控制器,负责在隧道端点设备上下发流表等数据面配置。其主要负责功能如下:

[0101] (1) 北向驱动层:负责监听存储系统中相应租户配置的变化,并进行后续处理。

[0102] (2) 北向逻辑处理层:用于提供参数校验、权限校验、第二控制器内部编排逻辑等功能。

[0103] (3) 北向数据模型和南向流表配置映射:负责将北向数据模型转换为南向流表配置;其中,北向数据模型为从存储系统获取到的数据所对应的模型,南向流表配置为需要下发的流表表项。

[0104] 例如,向流表配置包括对流表配置的定义、组织,并且,还可以对转换的流表进行持久化存储。此外,还可基于内部数据对账功能对北向数据和南向流表进行一致性校验,保证数据的准确性。

[0105] (4) 底层设备抽象层:用于对桥(bridge)、接口(interface)设备定义、校验处理、

以及持久化。

[0106] (5) 南向逻辑处理层:负责流标配置数据批量下发,还用于处理南向设备上报信息;该南向设备为下层的虚拟交换机等。

[0107] (6) 南向接口层:负责对接南向下发通道组件,例如对接下层的虚拟交换机。

[0108] 为便于描述,下面以一个实施例对整体的配置过程进行解释说明。其中,软件定义网络中包括多个VTEP设备,其作为隧道端点设备;两层控制器之间的存储系统采用ETCD。

[0109] 以图2所示的结构为例,若租户A创建了一台虚拟机1,这台虚拟机1被调度到了节点B,该节点B对应VTEP设备2;第一控制器记录各个VTEP设备下运行了哪个租户的虚拟机或容器,并根据租户在VTEP设备上创建的虚拟机或容器的资源情况,创建VTEP设备与租户之间的对应关系,并存储在ETCD中,其存储格式为:/<VTEP设备IP>/<租户ID>。

[0110] 例如,VTEP设备2的IP地址为1.1.1.1,租户A的ID为tenant\_A,则在ETCD中记录有:/1.1.1.1/tenant\_A,其为相应的索引结构。

[0111] 并且,第一控制器创建各个租户对应的逻辑配置信息,且均会关联相应的租户ID,该信息在ETCD中的存储格式为:/<逻辑配置类型>/<租户ID>/<配置项ID>。该存储格式具体为逻辑配置信息的存储路径,表示逻辑配置信息的键,其对应的值即为具体的配置内容。

[0112] 例如,该租户A对应有限速配置(rate-limit),其一个配置项ID为config001,则该逻辑配置信息的存储路径可以为:/rate-limit/tenant\_A/config001。

[0113] 对于各个VTEP设备,在第二控制器启动后,即可向ETCD查询本VTEP设备上所有租户信息。由于租户A的虚拟机不存在于VTEP设备1和VTEP设备3中,故这两个VTEP设备不会获取到关于租户A的任何配置。

[0114] 对于VTEP设备2,其第二控制器可以基于自身的IP地址向ETCD发起查询,通过前缀匹配查询的方式可以确定属于自身的租户标识tenant\_A。进一步地,VTEP设备2可以根据该租户标识tenant\_A组装ETCD前缀监听配置,进而监听到自身所需的逻辑配置信息。

[0115] 例如,可以组装出类似于“/\*/tenant\_A/”的前缀,其中的\*表示通配符,代表任何逻辑配置类型。通过这个前缀,VTEP设备2可以告诉ETCD,它对所有逻辑配置类型下,属于租户“tenant\_A”的所有配置变化都感兴趣。当ETCD中任何符合该前缀路径的配置发生变化时(无论是新增配置,修改配置,还是删除配置),ETCD都会向VTEP设备2的第二控制器发送通知,即监听事件。

[0116] VTEP设备2的第二控制器基于ETCD发送的配置变化的通知,即可从中获取具体的逻辑配置。例如,限速的具体数值、应用的网络接口等信息。这样,第二控制器能够及时获取到与自身对应的租户ID相关的逻辑配置变化,进而将这些变化下发给相关的网络设备,确保网络设备的流表与ETCD中存储的租户逻辑配置保持一致,实现对租户网络资源的准确管理。

[0117] 本实施例提供的软件定义网络的配置方法,软件定义网络采用基于索引的两层控制器分离架构,上层的第一控制器主要负责抽象和计算逻辑配置,对于集群内资源竞争的资源,因为第一控制器的单体计算能力也能得到很好的处理;隧道端点设备只获取有关的逻辑配置信息,即只负责自身设备下的资源管理,实现了网络配置按需下发,能够有效筛选和过滤无效的资源配置,避免将无用配置下发到数据面,有效降低了第二控制器和数据面的负载,有利于扩大整个集群的规模。

[0118] 在本实施例中还提供了一种软件定义网络的配置装置,该装置用于实现上述实施例及优选实施方式,已经进行过说明的不再赘述。如以下所使用的,术语“模块”可以是实现预定功能的软件和/或硬件的组合。尽管以下实施例所描述的装置较佳地以软件来实现,但是硬件,或者软件和硬件的组合的实现也是可能并被构想的。

[0119] 本实施例提供一种软件定义网络的配置装置,应用于软件定义网络的第一控制器,如图7所示,该装置包括:

[0120] 数据获取模块701,用于获取业务数据;所述业务数据包括各个租户的虚拟资源对应的逻辑配置信息以及所述虚拟资源对应的隧道端点设备;

[0121] 处理模块702,用于根据所述业务数据建立隧道端点设备与租户之间的对应关系,并确定与租户关联的逻辑配置信息;

[0122] 配置下发模块703,用于将所述隧道端点设备与租户之间的对应关系以及与租户关联的逻辑配置信息下发至下层的存储系统,以指示隧道端点设备从所述存储系统获取相应租户的逻辑配置信息。

[0123] 在一些可选的实施方式中,所述隧道端点设备与租户之间的对应关系所对应的索引结构,包括:第一层的隧道端点设备的设备标识,以及第二层的租户标识。

[0124] 在一些可选的实施方式中,所述处理模块702确定与租户关联的逻辑配置信息,包括:

[0125] 确定键值存储形式的与租户关联的逻辑配置信息;所述与租户关联的逻辑配置信息的键包括租户标识。

[0126] 本实施例提供一种软件定义网络的配置装置,应用于隧道端点设备的第二控制器,如图8所示,该装置包括:

[0127] 查询模块801,用于根据上层的存储系统中存储的隧道端点设备与租户之间的对应关系,查询与本地隧道端点设备相对应的目标租户;

[0128] 配置获取模块802,用于从所述存储系统获取与所述目标租户相关联的目标逻辑配置信息;所述存储系统存储有与租户关联的逻辑配置信息;

[0129] 流表下发模块803,用于根据所述目标逻辑配置信息生成相应的流表表项,并下发所述流表表项。

[0130] 在一些可选的实施方式中,所述隧道端点设备与租户之间的对应关系所对应的索引结构,包括:第一层的隧道端点设备的设备标识,以及第二层的租户标识;

[0131] 所述查询模块801根据上层的存储系统中存储的隧道端点设备与租户之间的对应关系,查询与本地隧道端点设备相对应的目标租户,包括:

[0132] 根据本地隧道端点设备的设备标识对存储系统中存储的隧道端点设备与租户之间的对应关系所对应的索引结构进行前缀匹配查询,确定目标租户的租户标识。

[0133] 在一些可选的实施方式中,所述存储系统包括存储和通知中间件;所述与租户关联的逻辑配置信息采用键值存储形式,且所述与租户关联的逻辑配置信息的键包括租户标识;

[0134] 所述配置获取模块802从所述存储系统获取与所述目标租户相关联的目标逻辑配置信息,包括:

[0135] 向所述存储和通知中间件发起监听请求;所述监听请求用于监听所述目标租户的

租户标识;

[0136] 获取在所述目标租户的租户标识所对应的值发生变化时,所述存储和通知中间件推送的监听事件;所述监听事件包括与所述目标租户相关联的目标逻辑配置信息。

[0137] 上述各个模块和单元的更进一步的功能描述与上述对应实施例相同,在此不再赘述。

[0138] 本实施例中的软件定义网络的配置装置是以功能单元的形式来呈现,这里的单元是指ASIC(Application Specific Integrated Circuit,专用集成电路)电路,包括执行一个或多个软件或固定程序的处理器和存储器,和/或其他可以提供上述功能的器件。

[0139] 本公开实施例还提供一种计算机设备,具有上述图7或图8所示的软件定义网络的配置装置。

[0140] 请参阅图9,图9是本公开可选实施例提供的一种计算机设备的结构示意图,如图9所示,该计算机设备包括:一个或多个处理器10、存储器20,以及用于连接各部件的接口,包括高速接口和低速接口。各个部件利用不同的总线互相通信连接,并且可以被安装在公共主板上或者根据需要以其它方式安装。处理器可以对在计算机设备内执行的指令进行处理,包括存储在存储器中或者存储器上以在外部输入/输出装置(诸如,耦合至接口的显示设备)上显示GUI的图形信息的指令。在一些可选的实施方式中,若需要,可以将多个处理器和/或多条总线与多个存储器一起使用。同样,可以连接多个计算机设备,各个设备提供部分必要的操作(例如,作为服务器阵列、一组刀片式服务器、或者多处理器系统)。图9中以一个处理器10为例。

[0141] 处理器10可以是中央处理器,网络处理器或其组合。其中,处理器10还可以进一步包括硬件芯片。上述硬件芯片可以是专用集成电路,可编程逻辑器件或其组合。上述可编程逻辑器件可以是复杂可编程逻辑器件,现场可编程逻辑门阵列,通用阵列逻辑或其任意组合。

[0142] 其中,所述存储器20存储有可由至少一个处理器10执行的指令,以使所述至少一个处理器10执行实现上述实施例示出的方法。

[0143] 存储器20可以包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需要的应用程序;存储数据区可存储根据计算机设备的使用所创建的数据等。此外,存储器20可以包括高速随机存取存储器,还可以包括非瞬时存储器,例如至少一个磁盘存储器件、闪存器件、或其他非瞬时固态存储器件。在一些可选的实施方式中,存储器20可选包括相对于处理器10远程设置的存储器,这些远程存储器可以通过网络连接至该计算机设备。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。

[0144] 存储器20可以包括易失性存储器,例如,随机存取存储器;存储器也可以包括非易失性存储器,例如,快闪存储器,硬盘或固态硬盘;存储器20还可以包括上述种类的存储器的组合。

[0145] 该计算机设备还包括通信接口30,用于该计算机设备与其他设备或通信网络通信。

[0146] 本公开实施例还提供了一种计算机可读存储介质,上述根据本公开实施例的方法可在硬件、固件中实现,或者被实现为可记录在存储介质,或者被实现通过网络下载的原始

存储在远程存储介质或非暂时机器可读存储介质中并将被存储在本地存储介质中的计算机代码,从而在此描述的方法可被存储在使用通用计算机、专用处理器或者可编程或专用硬件的存储介质上的这样的软件处理。其中,存储介质可为磁碟、光盘、只读存储记忆体、随机存储记忆体、快闪存储器、硬盘或固态硬盘等;进一步地,存储介质还可以包括上述种类的存储器的组合。可以理解,计算机、处理器、微处理器控制器或可编程硬件包括可存储或接收软件或计算机代码的存储组件,当软件或计算机代码被计算机、处理器或硬件访问且执行时,实现上述实施例示出的方法。

[0147] 本公开的一部分可被应用为计算机程序产品,例如计算机程序指令,当其被计算机执行时,通过该计算机的操作,可以调用或提供根据本公开的方法和/或技术方案。本领域技术人员应能理解,计算机程序指令在计算机可读介质中的存在形式包括但不限于源文件、可执行文件、安装包文件等,相应地,计算机程序指令被计算机执行的方式包括但不限于:该计算机直接执行该指令,或者该计算机编译该指令后再执行对应的编译后程序,或者该计算机读取并执行该指令,或者该计算机读取并安装该指令后再执行对应的安装后程序。在此,计算机可读介质可以是可供计算机访问的任意可用的计算机可读存储介质或通信介质。

[0148] 虽然结合附图描述了本公开的实施例,但是本领域技术人员可以在不脱离本公开的精神和范围的情况下做出各种修改和变型,这样的修改和变型均应涵盖在本公开的保护范围之内。

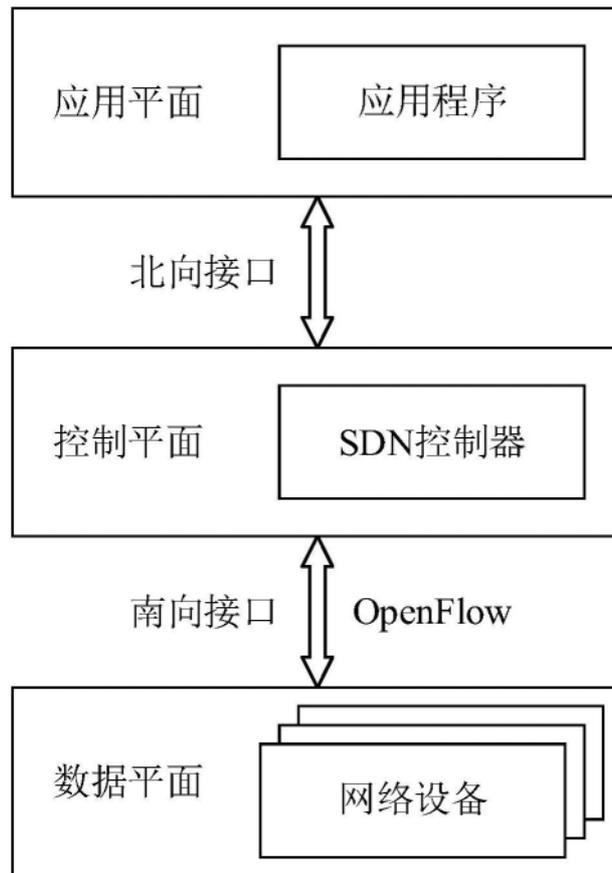


图1

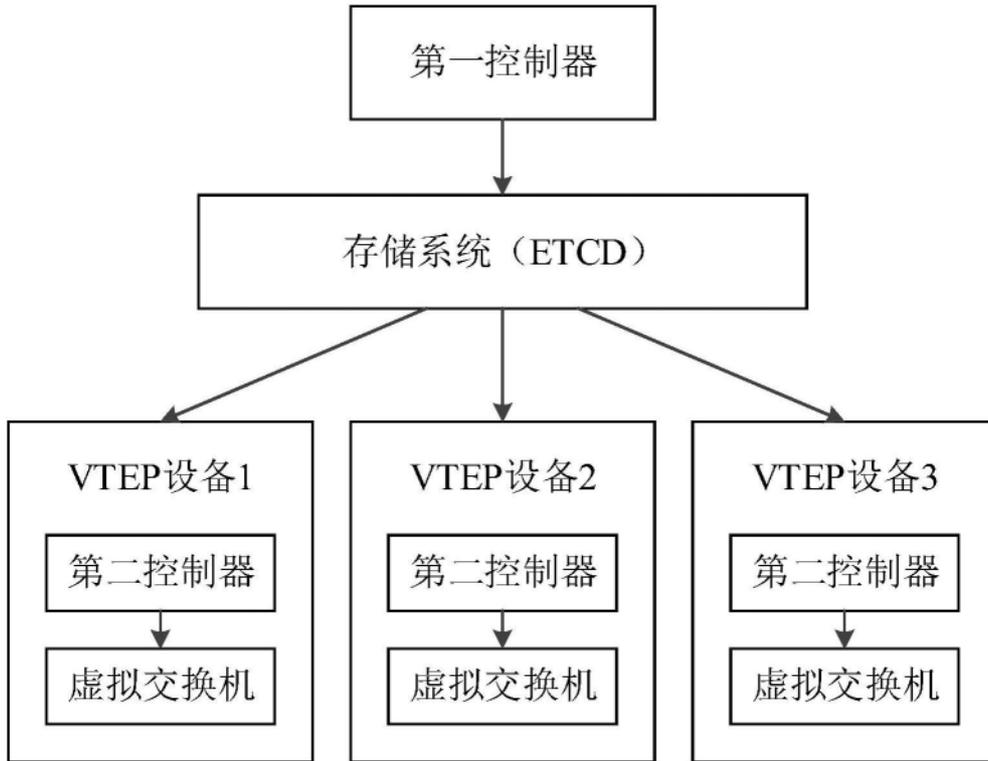


图2

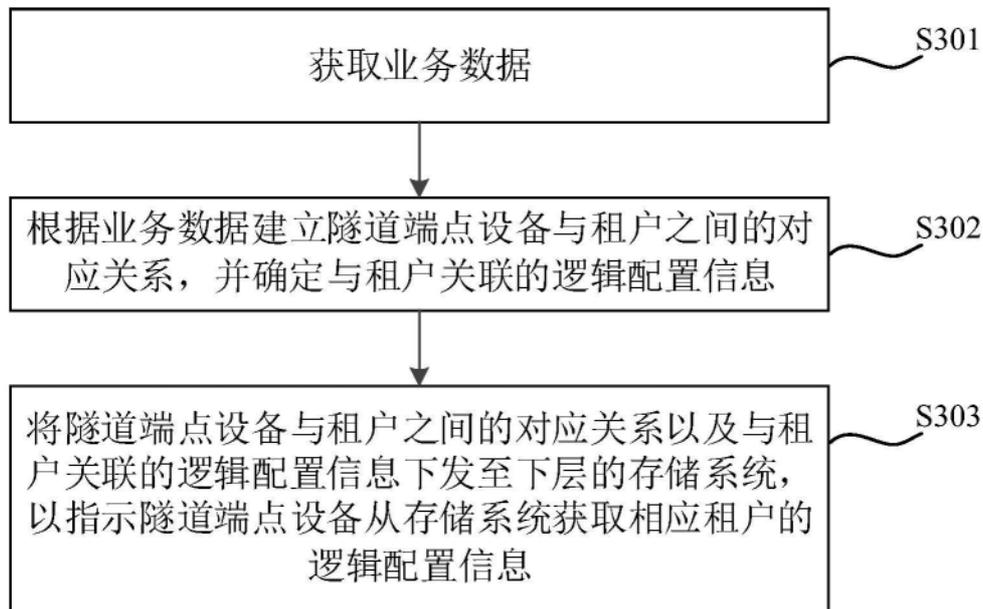


图3

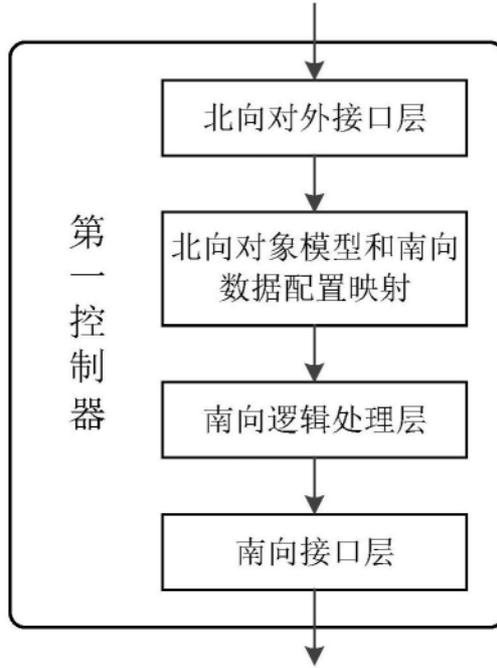


图4

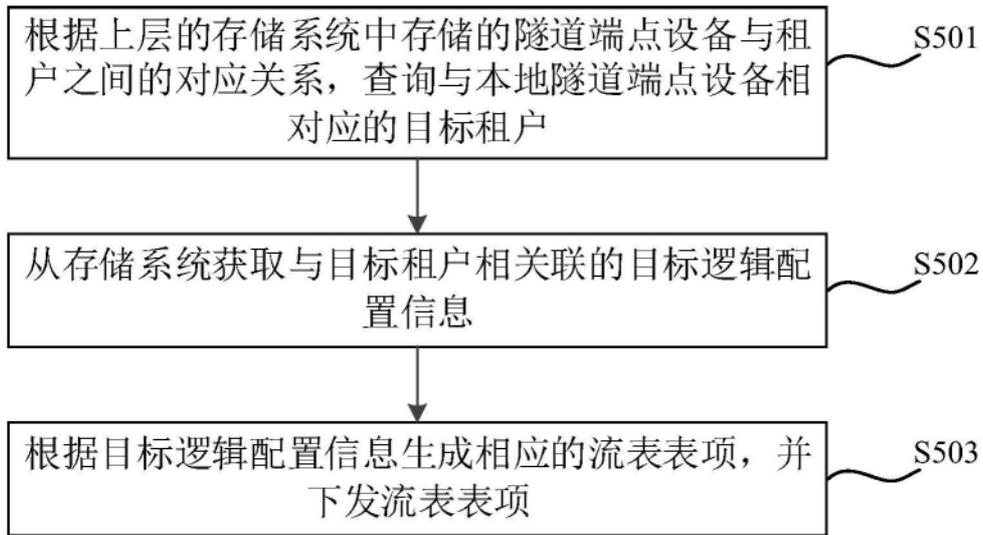


图5

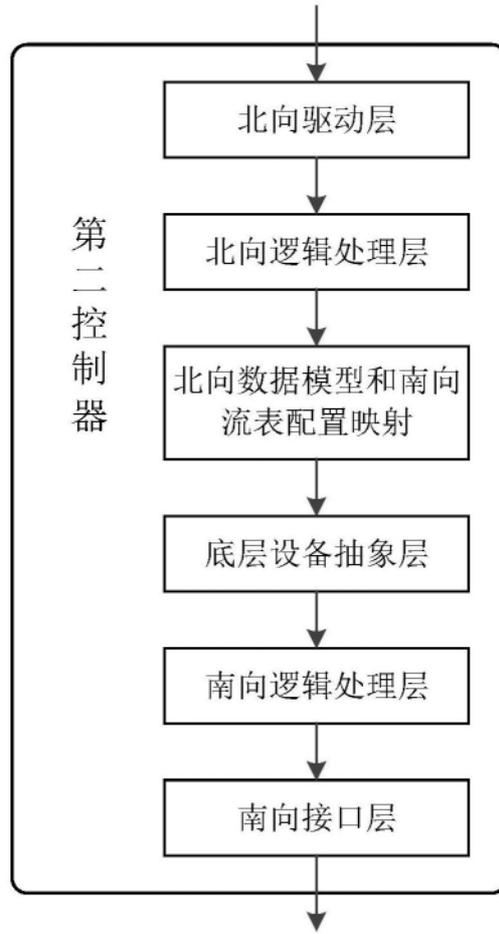


图6

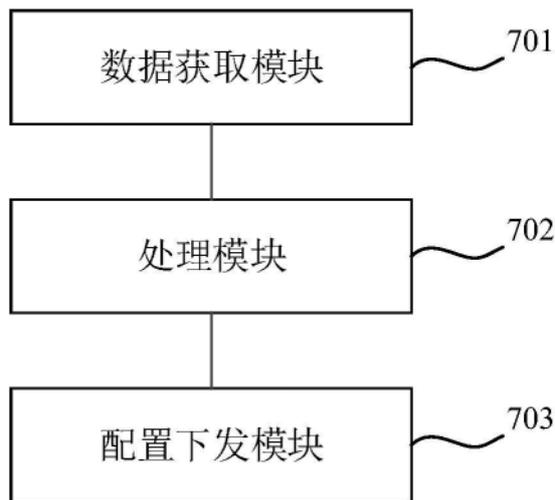


图7

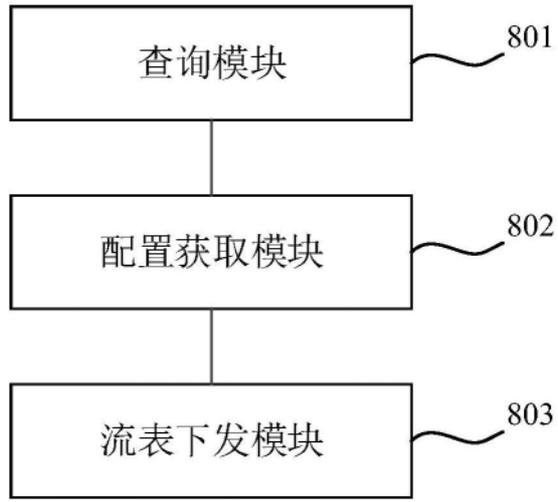


图8

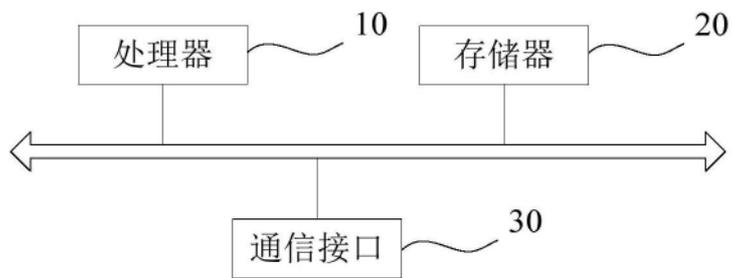


图9