



(12)发明专利

(10)授权公告号 CN 106101280 B

(45)授权公告日 2019.01.22

(21)申请号 201610685135.5

(22)申请日 2016.08.18

(65)同一申请的已公布的文献号
申请公布号 CN 106101280 A

(43)申请公布日 2016.11.09

(73)专利权人 无锡华云数据技术服务有限公司
地址 214000 江苏省无锡市滨湖区科教软
件园6号

(72)发明人 许广彬 郑军 张银滨 刘继江
何少鹏 谢旻 田大鹏

(74)专利代理机构 北京商专永信知识产权代理
事务所(普通合伙) 11400
代理人 高之波 储振

(51)Int.Cl.
H04L 29/08(2006.01)

(56)对比文件

CN 104243527 A,2014.12.24,
CN 103581177 A,2014.02.12,
CN 104519130 A,2015.04.15,
CN 105763439 A,2016.07.13,
EP 2767051 A1,2014.08.20,
US 7609619 B2,2009.10.27,
US 2014149552 A1,2014.05.29,
WO 2014086023 A1,2014.06.12,
Wenzhi Cui等.DiFS: distributed flow
scheduling for adaptive routing in
hierarchical data center networks.《ANCS ”
14 Proceedings of the tenth ACM/IEEE
symposium on Architectures for networking
and communications systems》.2014,P53-64.

审查员 张浩

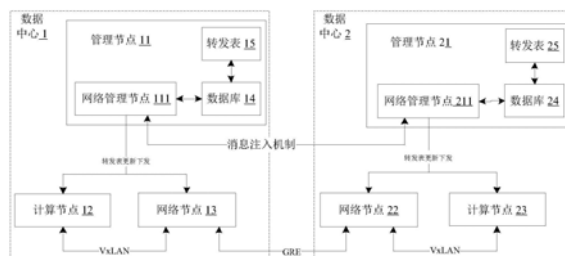
权利要求书1页 说明书4页 附图1页

(54)发明名称

一种数据中心间的网络信息同步更新方法

(57)摘要

本发明提供了数据中心间的网络信息同步更新方法,包括以下步骤:S1、多个数据中心之间通过其所属的网络节点基于GRE实现互联;S2、对发生更新的管理节点通过远程过程调用消息注入机制对其他数据中心的网络节点中的转发表进行更新操作;S3、通过管理节点将更新后的转发表对每个数据中心所属的网络节点和/或计算节点进行更新操作。在本发明中,通过将网络信息的同步操作从数据层面转移到控制层面,避免了使用手动配置或者BGP协议对网络信息的同步更新,降低了多个数据中心之间在网络信息的同步更新过程中因同步更新操作对网络带宽造成开销,提高数据中心间网络信息的数据一致性,降低了对硬件设备的依赖性。



1. 数据中心间的网络信息同步更新方法,其特征在于,包括以下步骤:

S1、多个数据中心之间通过其所属的网络节点基于通用路由封装协议实现互联;

S2、当某一数据中心中的管理节点中的转发表发生更新时,发生更新的管理节点通过远程过程调用消息注入机制对其他数据中心的管理节点中的转发表进行更新操作;

S3、通过管理节点使用更新后的转发表对每个数据中心所属的网络节点和/或计算节点进行更新操作。

2. 根据权利要求1所述的数据中心间的网络信息同步更新方法,其特征在于,所述转发表存储于管理节点的数据库中。

3. 根据权利要求1所述的数据中心间的网络信息同步更新方法,其特征在于,所述管理节点被配置为物理机、虚拟机、物理集群服务器或者虚拟机集群服务器。

4. 根据权利要求1所述的数据中心间的网络信息同步更新方法,其特征在于,所述步骤S3具体为:通过管理节点将更新后的转发表通过二层网络消息注入机制分别对每个数据中心所属的网络节点和/或计算节点进行更新操作。

一种数据中心间的网络信息同步更新方法

技术领域

[0001] 本发明涉及云计算技术领域,尤其涉及一种数据中心间的网络信息同步更新方法。

背景技术

[0002] 在云计算数据中心中,数据中心之间大二层网络需要保证跨数据中心的虚拟机之间业务流量能够像数据中心内部本地的流量一样传输。一般情况下,数据中心之间的互联需要穿越运营商的三层网络,数据中心之间使用L2over L3的方式通过VxLAN或者GRE隧道进行互联。数据中心之间通过隧道进行二层互通的关键点是本地数据中心需要知道对端数据中心虚拟机的路由信息,通常的做法是人工在各个数据中心的边缘节点上手工配置这些信息,或者通过BGP路由的方式在每个租户的隧道内部相互通告这些信息,手工方式不能实现自动化配置,并且数据中心内部租户和虚拟机众多,配置繁琐容易出错,BGP的方式需要在数据中心的边缘节点部署BGP路由器,BGP路由实现复杂,并且BGP通告会占用数据链路的带宽。

发明内容

[0003] 本发明的目的在于公开一种数据中心间的网络信息同步更新方法,用以在多个数据中心之间在实现互联过程中实现网络信息的同步更新,通过避免使用手动配置或者BGP协议对网络信息的同步更新,同时降低多个数据中心之间在网络信息的同步更新过程中因同步更新操作对网络带宽造成开销,提高多个数据中心之间数据的一致性,并降低网络信息更新对硬件设备的依赖性。

[0004] 为实现上述目的,本发明提供了一种数据中心间的网络信息同步更新方法,包括以下步骤:

[0005] S1、多个数据中心之间通过其所属的网络节点基于GRE实现互联;

[0006] S2、当某一数据中心中的管理节点中的转发表发生更新时,对发生更新的管理节点通过远程过程调用消息注入机制对其他数据中心的管理节点中的转发表进行更新操作;

[0007] S3、通过管理节点将更新后的转发表对每个数据中心所属的网络节点和/或计算节点进行更新操作。

[0008] 作为本发明的进一步改进,所述转发表存储于管理节点的数据库中。

[0009] 作为本发明的进一步改进,所述管理节点被配置为物理机、虚拟机、物理集群服务器或者虚拟机集群服务器。

[0010] 作为本发明的进一步改进,所述步骤S3具体为:通过管理节点将更新后的转发表通过二层网络消息注入机制分别对每个数据中心所属的网络节点和/或计算节点进行更新操作。

[0011] 与现有技术相比,本发明的有益效果是:在本发明中,转发表可在各个数据中心中的管理节点之间的更新,并将更新后的转发表对每个数据中心中的网络节点和/或计算节

点进行更新操作,因此通过将网络信息的同步操作从数据层面转移到控制层面,避免了使用手动配置或者BGP协议对网络信息的同步更新,降低了多个数据中心之间在网络信息的同步更新过程中因同步更新操作对网络带宽造成开销,提高多个数据中心之间网络信息的数据一致性,显著地降低了因网络信息更新对硬件设备的依赖性。

附图说明

[0012] 图1为本发明一种数据中心间的网络信息同步更新方法的流程图;

[0013] 图2为示出了两个数据中心互联时的拓扑图。

具体实施方式

[0014] 下面结合附图所示的各实施方式对本发明进行详细说明,但应当说明的是,这些实施方式并非对本发明的限制,本领域普通技术人员根据这些实施方式所作的功能、方法、或者结构上的等效变换或替代,均属于本发明的保护范围之内。

[0015] 在详细描述本发明之前,首先对说明书中涉及到几个主要技术术语进行简要阐述。

[0016] 边界网关协议(BGP),运行于TCP上的一种自治系统的路由协议。BGP是用来处理像因特网大小的网络的协议,其能够妥善处理好不相关路由域间的多路连接的协议。BGP构建在EGP的经验之上。BGP系统的主要功能是和其他的BGP系统交换网络可达信息。网络可达信息包括列出的自治系统(AS)的信息。这些信息有效地构造了AS互联的拓扑图,并由此清除了路由环路,同时在AS级别上可实施策略决策。

[0017] 通用路由封装协议(GRE),一组路由器在统一管理之下,在AS内使用内部网关协议和统一度量来路由数据包,并通过外部网关协议将数据包路由到其他AS中。GRE可对某些网络层协议(如IP和IPX)的数据报文进行封装,使这些被封装的数据报文能够在另一个网络层协议(如IP)中传输。GRE采用了Tunnel(隧道)技术,是VPN(Virtual Private Network)的第三层隧道协议。

[0018] VxLAN(Virtual Extensible Local Area Network)是一种将二层报文用三层协议进行封装的技术,可以对二层网络在三层范围进行扩展。它应用于数据中心内部,使虚拟机可以在互相连通的三层网络范围内迁移,而不需要改变IP地址和MAC地址,保证业务的连续性。VxLAN采用24bit的网络标识,使用户可以创建16M相互隔离的虚拟网络,突破了目前广泛采用的VLAN所能表示的4K个隔离网络的应用限制,这使得大规模多租户的云环境中具有了充足的虚拟网络分区资源。

[0019] 请参阅图1至图2所示出的本发明一种数据中心间的网络信息同步更新方法的一种具体实施方式。

[0020] 在本实施方式,该数据中心间的网络信息同步更新方法,包括以下步骤:

[0021] 步骤S1、多个数据中心之间通过其所属的网络节点基于GRE实现互联。

[0022] 步骤S2、当某一数据中心中的管理节点中的转发表发生更新时,对发生更新的管理节点通过远程过程调用消息注入机制对其他数据中心的管理节点中的转发表进行更新操作。通过该步骤S2,可实现当某一个数据中心中的管理节点中的转发表发生更新时,在不同物理位置中的数据中心(甚至可以是物理机)上的进程通过远程调用的方式相互通行。

[0023] 步骤S3、通过管理节点将更新后的转发表对每个数据中心所属的网络节点和/或计算节点进行更新操作。

[0024] 其中,所述转发表存储于管理节点的数据库中。管理节点被配置为物理机、虚拟机、物理集群服务器或者虚拟机集群服务器。每个数据中心中的网络节点与计算节点之间通过基于VxLAN实现互联。在本实施方式中,该步骤S3具体为:通过管理节点将更新后的转发表通过二层网络消息注入机制分别对每个数据中心所属的网络节点和/或计算节点进行更新操作。

[0025] 在图1中,包含数据中心1与数据中心2为例示范性说明,更多的数据中心可以用更多的数据中心之间互连叠加实现。在图1中,数据中心1及数据中心2均具有独立的云平台系统,例如openstack系统。在本实施方式中,为便于描述,对基于openstack系统的数据中心实现网络信息的同步更新的过程进行详细描述。

[0026] 数据中心1中包括管理节点11,管理节点11中包含一个网络管理节点111、与网络管理节点111互联的数据库14,存储于数据库14中的转发表15,计算节点12以及网络节点13。管理节点11为数据中心1的控制平面。计算节点12与网络节点13共同构成数据中心1的数据平面。数据中心2中包括管理节点21,管理节点21中包含一个网络管理节点211、与网络管理节点211互联的数据库24,存储于数据库24中的转发表25,计算节点23以及网络节点22。管理节点21为数据中心2的控制平面。计算节点23与网络节点22共同构成数据中心2的数据平面。其中,计算节点23与网络节点22通过VxLAN互联,计算节点12与网络节点13通过VxLAN互联。转发表用于记录用户在不同数据中心进行网络访问时的参数,转发表中分别形成若干条记录,每一条记录由用户id、用户ip及用户mac地址构成。

[0027] 在本实施方式中,数据中心1的管理节点11与数据中心2中的管理节点21相互独立,并且没有采用级联或者Cell等方案相互关联。但基于这些openstack系统,可能会有用户的上层管理程序,从而保证同一租户在不同的数据中心创建的虚拟机具有相同的租户id。如果没有这种上层管理程序,openstack管理员就要人工保持同一租户的id一致性了。

[0028] 在openstack网络节点中,数据中心1与数据中心2通过分别所属的网络节点13及网络节点22基于GRE创建一个跨数据中心的连接,以形成一个跨数据中心的大二层网络。数据中心1的网络节点13与数据中心2的网络节点22建立连接之后,管理节点11就可以和对端的管理节点21互相通告转发表。

[0029] 管理节点11与管理节点21中各自生成并独立维护一个转发表15及转发表25。转发表15记录有数据中心1中每个计算机的转发信息。在本实施方式中,该计算机应当作广义理解,包括但不限于台式机、笔记本电脑、虚拟机、服务器、手机或者其他具有数据处理功能的数字设备。同时,转发表25记录有数据中心2中每个计算机的转发信息。

[0030] 当数据中心1的网络连接发生变化(例如虚拟机创建操作,虚拟机迁移操作,虚拟机删除操作等),数据中心1中管理节点11所属的网络管理节点111需要对存储于数据库14中的转发表15进行更新操作,然后将更新后的转发表通过消息注入机制通知数据中心2中的管理节点21所属的网络管理节点211,以通过网络管理节点211对与其互联的数据库24中原来存储的转发表25进行更新操作。在此过程中,转发表的更新操作发生在数据中心1与数据中心2的控制平面。因此,消除了该转发表在执行更新操作时对网络节点13及网络节点22的带宽资源的占用,降低了多个数据中心之间在网络信息的同步更新过程中因同步更新操

作对网络带宽造成开销。

[0031] 当数据中心2中的管理节点21完成了转发表的更新操作后,网络管理节点211可对网络节点22或者计算节点23或者同时对网络节点22及计算节点23进行转发表更新下发操作,以将数据中心1中已经发生变化的转发表通知数据中心2中的网络节点22和/或计算节点23,以实现位于数据层面中的网络节点与计算节点的同步通知,避免了位于不同数据中心中的数据层面上的计算节点与网络节点对网络信息的学习或者配置。

[0032] 当数据中心2接到更新后的转发表后,通过二层信息分发机制下发到这次更新会影响到所有计算节点22和网络节点23,从而使得数据中心1的发生更改的网络信息同步到数据中心2中。例如,当数据中心2被配置为开源平台(例如openstack平台)时,该二层信息分发机制为L2population。进一步的,该数据中心2或者数据中心1也被配置为非开源平台(例如VMware公司的VSPHERE),并通过远程过程调用(RPC)下发到这次更新会影响到所有计算节点22和网络节点23,从而使得数据中心1的发生更改的网络信息同步到数据中心2中。

[0033] 当然,如果数据中心1中转发表15发生变化时,网络管理节点111也需要通过二层信息分发机制下发到这次更新会影响到所有计算节点12和网络节点13。

[0034] 数据中心2中的转发表25发生变化并对数据中心1中的转发表15进行同步更新的具体过程与数据中心1中的转发表15对数据中心2中的转发表25进行同步更新的具体过程相同,具体请参上文所述,在此不再赘述。

[0035] 上文所列出的一系列详细说明仅仅是针对本发明的可行性实施方式的具体说明,它们并非用以限制本发明的保护范围,凡未脱离本发明技艺精神所作的等效实施方式或变更均应包含在本发明的保护范围之内。

[0036] 对于本领域技术人员而言,显然本发明不限于上述示范性实施例的细节,而且在不背离本发明的精神或基本特征的情况下,能够以其他的具体形式实现本发明。因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权利要求要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化囊括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。

[0037] 此外,应当理解,虽然本说明书按照实施方式加以描述,但并非每个实施方式仅包含一个独立的技术方案,说明书的这种叙述方式仅仅是为清楚起见,本领域技术人员应当将说明书作为一个整体,各实施例中的技术方案也可以经适当组合,形成本领域技术人员可以理解的其他实施方式。

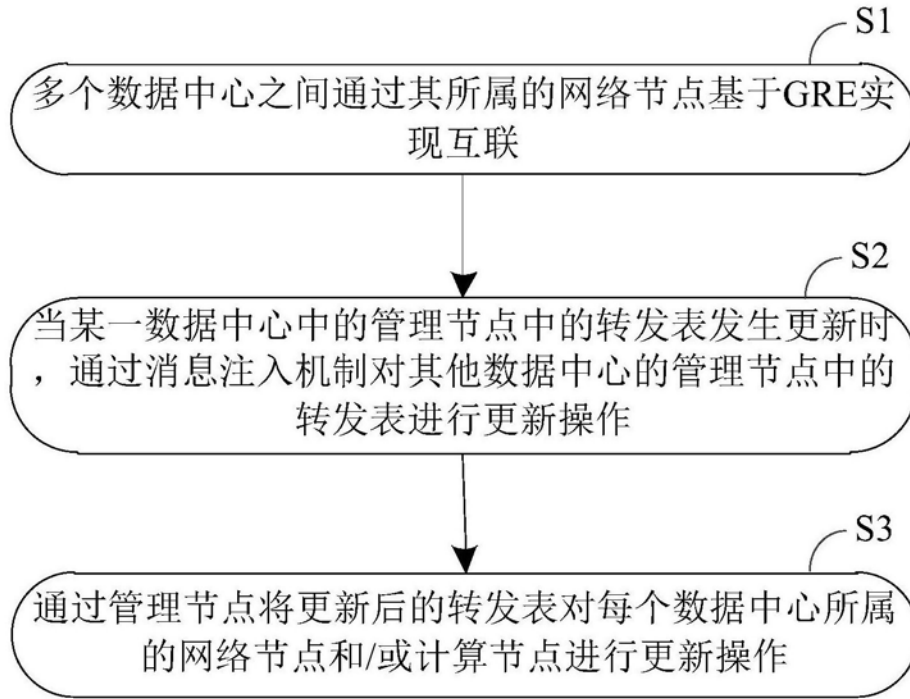


图1

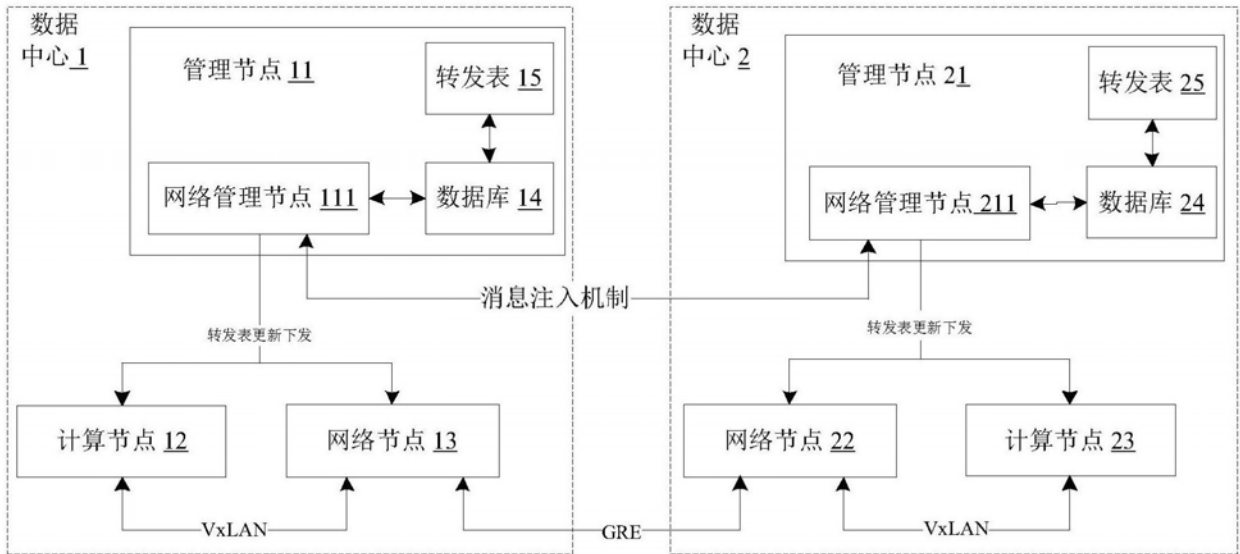


图2