(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2016 年 3 月 10 日 (10.03.2016)

WIPO | PCT

(10) 国际公布号
WO 2016/033785 A1

(51) 国际专利分类号:
H04L 12/757 (2013.01)

(21) 国际申请号: PCT/CN2014/085944

(22) 国际申请日: 2014 年 9 月 4 日 (04.09.2014)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: **华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.)** [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(72) 发明人: **程璞 (CHENG, Pu)**; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 **谢银桥 (XIE, Yinqiao)**; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(81) 指定国 (除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明，要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

(54) **Title**: METHOD FOR SYNCHRONIZING FORWARDING TABLES, NETWORK DEVICE, AND SYSTEM

(54) **发明名称** : 转发表同步的方法、网络设备和系统

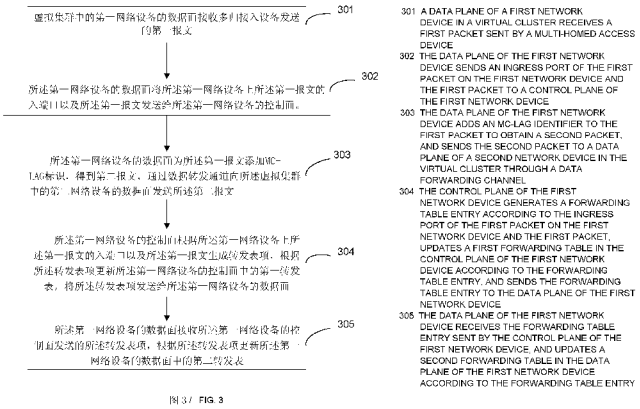


图 3 / FIG. 3

(57) **Abstract**: The present invention provides a method for synchronizing forwarding tables, a network device, and a system. A data plane of a first network device receives a first packet sent by a multi-homed access device; the data plane of the first network device sends the first packet to a control plane of the first network device; the data plane of the first network device adds an MC-LAG identifier to the first packet to obtain a second packet, and sends the second packet to a data plane of a second network device through a data forwarding channel; the control plane of the first network device generates a forwarding table entry according to an ingress port of the first packet on the first network device and the first packet, updates a first forwarding table in the control plane of the first network device according to the forwarding table entry, and sends the forwarding table entry to the data plane of the first network device; and the data plane of the first network device updates a second forwarding table in the data plane of the first network device according to the forwarding table entry. By means of the present invention, forwarding tables of network devices in a virtual cluster can be synchronized.

(57) **摘要**:

[见续页]

WO 2016/033785 A1

本发明提供一种转发表同步的方法、网络设备和系统。第一网络设备的数据面接收多归接入设备发送的第一报文；将所述第一报文发送给所述第一网络设备的控制面；为所述第一报文添加 MC-LAG 标识，得到第二报文，通过数据转发通道向所述第二网络设备的数据面发送所述第二报文；所述第一网络设备的控制面根据所述第一网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述第一网络设备的控制面中的第一转发表，将所述转发表项发送给所述第一网络设备的数据面；所述第一网络设备的数据面根据所述转发表项更新所述第一网络设备的数据面中的第二转发表。通过本发明，可以实现虚拟集群中各网络设备的转发表的同步。

转发表同步的方法、网络设备和系统

技术领域

本发明涉及通信领域，尤其涉及一种转发表同步的方法、网络设备和系统。

背景技术

链路聚合组（英文：Link Aggregation Group,简称：LAG）是一种基于多条以太链路的反向复用（英文：inverse multiplexing）方法，将多条连接到同一设备的链路捆绑在一起，能够增加链路带宽并提供冗余备份，提高链路的可靠性。链路聚合组内的多条链路可以当作一条逻辑链路。在以太网中，链路和端口是一一对应的，因此，链路聚合又被称为端口聚合。链路聚合组中的所有端口共用同一个 MAC 地址。

多机框链路聚合组（英文：Multi-Chassis Link Aggregation Group，简称：MC-LAG）是在 LAG 提供的链路级冗余之外增加的设备级冗余，其允许多个网络设备连接到共同的多归接入设备。对于所述多归接入设备来说，所述多个网络设备表现为一个逻辑 LAG。其中，一个 MAC-LAG 中包括的多个网络设备可以是交换机，路由器等。所述多归接入设备可以是交换机，路由器等网络设备，也可以是服务器，主机等非网络设备。

如图 1 所示，为 MC-LAG 的结构示意图。多归接入设备通过虚拟集群访问因特网。该虚拟集群中至少包括两个网络设备，图 1 中示出了网络设备 A 和网络设备 B，所述至少两个网络设备分别连接所述多归接入设备的链路组成了一个 MC-LAG。从所述多归接入设备的角度，只能看到一条链路，即该 MC-LAG。所述至少两个网络设备既能实现负载分担又能实现相互备份。所述负载分担是指，所述至少两个网络设备分别转发根据一定的规则确定的所述多归接入设备的部分流量；所述相互备份是指，所述至少两个网络设备中任意一个发生故障时，所述故障网络设备上的流量可以实时切换到另一个没有发生故障的网络设备上，保证用户业务不中断。

为了对业务流进行转发，需要在网络设备上生成用于转发业务流的转发表，如图 1 所示，所述转发表可以包括媒体访问控制（英文：Media Access Control，简称:MAC）地址表，地址解析协议（英文：Address Resolution Protocol，简称：ARP）表，组播表，路由表中的至少一种。此外，还可以包括其他转发表。为了在所述至少两个网络设备上实现负载分担以及相互备份，所述至少两个网络设备上的转发表必须保持一致，所述一致，是指所述两个网络设备上的转发表都必须包括所述多归接入设备的每个业务流的转发表项。因此，所述至少两个网络设备之间需要同步转发表。当网络设备的控制面

和数据面分离的时候，每个网络设备的控制面的转发表和数据面的转发表需要保持一致。

现有技术中为了实现转发表的一致性，采用的方法是，所述至少两个网络设备的控制面之间同步转发表。然后，所述至少两个网络设备的控制面分别将同步后的转发表发送给各自的数据面。

发明人在研究过程中发现，上述方法需要虚拟集群中的至少两个网络设备的控制面之间定时或者实时同步转发表，而控制面之间发送的是控制协议报文，有软件兼容性要求，即要求所述至少两个网络设备能够识别同样的控制协议，因此，当一个网络设备的软件版本升级之后，需要相应升级其他网络设备的软件版本，实现复杂。并且，当所述至少两个网络设备没有同时升级软件版本时，会导致所述至少两个网络设备上的转发表不一致，进而导致流量转发错误，用户业务中断。

发明内容

本发明用于提供一种转发表同步的方法、网络设备和系统，用于解决现有技术中，虚拟集群中的网络设备没有同时升级软件版本，引起的转发表不一致的问题。

本发明第一方面提供一种转发表同步的方法，应用于虚拟集群中，所述虚拟集群至少包括第一网络设备和第二网络设备，所述第一网络设备通过数据转发通道与所述第二网络设备通信，所述第一网络设备和第二网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组MC-LAG；所述方法包括：

所述第一网络设备的数据面接收所述多归接入设备发送的第一报文；将所述第一网络设备上所述第一报文的入端口所述第一报文发送给所述第一网络设备的控制面；为所述第一报文添加MC-LAG标识，得到第二报文，通过所述数据转发通道向所述第二网络设备的数据面发送所述第二报文；所述MC-LAG标识用于指示所述第一网络设备上所述第一报文的入端口为所述MC-LAG的成员端口；

所述第一网络设备的控制面根据所述第一网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述第一网络设备的控制面中的第一转发表，将所述转发表项发送给所述第一网络设备的数据面；

所述第一网络设备的数据面接收所述第一网络设备的控制面发送的所述转发表项，根据所述转发表项更新所述第一网络设备的数据面中的第二转发表。

结合所述第一方面，在所述第一方面的第一种实施方式中，所述第一网络设备的数

据面为所述第一报文添加MC-LAG标识包括：

所述第一网络设备的数据面根据所述第一网络设备上所述第一报文的入端口以及预先配置的包括MC-LAG标识与端口的对应关系的MC-LAG表，获取所述入端口对应的MC-LAG标识，为所述第一报文添加所述获取的MC-LAG标识。

结合所述第一方面或者第一方面的第一种实施方式，在所述第一方面的第二种实施方式中，所述第一网络设备的数据面将所述第一网络设备上所述第一报文的入端口以及所述第一报文发送给所述第一网络设备的控制面包括：

所述第一网络设备的数据面判断所述第一报文是否匹配所述第一网络设备的控制面下发的报文选择规则，如果所述第一报文匹配所述控制面下发的报文选择规则，将所述第一网络设备上所述第一报文的入端口以及所述第一报文发送给所述第一网络设备的控制面。

本发明第二方面提供另一种转发表同步的方法，应用于虚拟集群中，所述虚拟集群至少包括第一网络设备和第二网络设备，所述第一网络设备通过数据转发通道与所述第二网络设备通信，所述第一网络设备和第二网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组MC-LAG；所述方法包括：

所述第二网络设备的数据面通过所述数据转发通道接收所述第一网络设备的数据面发送的第二报文，所述第二报文是由所述第一网络设备为所述多归接入设备发送的第一报文添加MC-LAG标识后生成的，所述MC-LAG标识用于指示所述第一网络设备上所述第一报文的入端口为所述MC-LAG的成员端口；

所述第二网络设备的数据面根据所述第二报文，得到所述第一报文；

所述第二网络设备的数据面确定所述第二网络上所述第一报文的入端口，并将所述确定的入端口和所述第一报文发送给所述第二网络设备的控制面；

所述第二网络设备的控制面根据所述确定的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述控制面中的第一转发表，并将所述转发表项发送给所述第二网络设备的数据面；

所述第二网络设备的数据面接收所述第二网络设备的控制面发送的所述转发表项，根据所述转发表项更新所述第二网络设备的数据面中的第二转发表。

结合所述第二方面，在所述第二方面的第一种实施方式中，所述根据所述第二报文，得到所述第一报文包括：

所述第二网络设备的数据面从所述第二报文中删除所述MC-LAG标识，得到所述第一

报文。

结合所述第二方面或者第二方面的第一种实施方式，在所述第二方面的第二种实施方式中，所述确定所述第二网络设备上所述第一报文的入端口包括：

所述第二网络设备的数据面根据所述第二报文获取所述MC-LAG标识，根据预先配置的包括MC-LAG标识与端口的对应关系的MC-LAG表，获取对应所述MC-LAG标识且属于所述第二网络设备的端口，将所述获取到的端口作为所述第二网络设备上所述第一报文的入端口。

本发明第三方面提供一种网络设备，应用于虚拟集群中，所述虚拟集群至少包括所述网络设备和另一网络设备，所述网络设备通过数据转发通道与所述另一网络设备通信，所述网络设备和另一网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组MC-LAG；

所述网络设备包括控制面和数据面；

所述数据面用于接收所述多归接入设备发送的第一报文；将所述网络设备上所述第一报文的入端口以及所述第一报文发送给所述控制面；为所述第一报文添加MC-LAG标识，得到第二报文，通过所述数据转发通道向所述另一网络设备的数据面发送所述第二报文；所述MC-LAG标识用于指示所述网络设备上所述第一报文的入端口为所述MC-LAG的成员端口；

所述控制面用于根据所述网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述控制面中的第一转发表，将所述转发表项发送给所述数据面；

所述数据面还用于接收所述控制面发送的所述转发表项，根据所述转发表项更新所述数据面中的第二转发表。

结合所述第三方面，在所述第三方面的第一种实施方式中，为所述第一报文添加MC-LAG标识时，所述数据面具体用于：

根据所述网络设备上所述第一报文的入端口以及预先配置的包括MC-LAG标识与端口的对应关系的MC-LAG表，获取所述入端口对应的MC-LAG标识，为所述第一报文添加所述获取的MC-LAG标识。

结合所述第三方面或者第三方面的第一种实施方式，在所述第三方面的第二种实施方式中，将所述网络设备上所述第一报文的入端口以及所述第一报文发送给所述控制面时，所述数据面具体用于：

判断所述第一报文是否匹配所述控制面下发的报文选择规则，如果所述第一报文匹配所述控制面下发的报文选择规则，将所述网络设备上所述第一报文的入端口以及所述第一报文发送给所述控制面。

本发明第四方面提供另一种网络设备，应用于虚拟集群中，所述虚拟集群至少包括所述网络设备和另一网络设备，所述网络设备通过数据转发通道与所述另一网络设备通信，所述网络设备和另一网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组MC-LAG；

所述网络设备包括控制面和数据面；

所述数据面通过所述数据转发通道接收所述另一网络设备的数据面发送的第二报文，所述第二报文是由所述另一网络设备为所述多归接入设备发送的第一报文添加MC-LAG标识后生成的，所述MC-LAG标识用于指示所述另一网络设备上所述第一报文的入端口为所述MC-LAG的成员端口；根据所述第二报文，得到所述第一报文；确定所述网络设备上所述第一报文的入端口，并将所述确定的入端口和所述第一报文发送给所述控制面；

所述控制面用于根据所述确定的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述控制面中的第一转发表，并将所述转发表项发送给所述数据面；

所述数据面还用于接收所述控制面发送的所述转发表项，根据所述转发表项更新所述数据面中的第二转发表。

结合所述第四方面，在所述第四方面的第一种实施方式中，在根据所述第二报文，得到所述第一报文时，所述数据面具体用于从所述第二报文中删除所述MC-LAG标识，得到所述第一报文。

结合所述第四方面或者第四方面的第一种实施方式，在所述第四方面的第二种实施方式中，在确定所述网络设备上所述第一报文的入端口时，所述数据面具体用于：

根据所述第二报文获取所述MC-LAG标识，根据预先配置的包括MC-LAG标识与端口的对应关系的MC-LAG表，获取对应所述MC-LAG标识且属于所述网络设备的端口，将所述获取到的端口作为所述网络设备上所述第一报文的入端口。

本发明第五方面提供一种转发表同步的系统，所述系统包括第一网络设备和第二网络设备；所述第一网络设备为所述第三方面、第三方面的第一种实施方式或第二种实施方式中的网络设备；所述第二网络设备为所述第四方面、第四方面的第一种实施方式或第二种实施方式中的网络设备。

本发明中，虚拟集群中的第一网络设备的数据面收到多归接入设备发送的第一报文后，将该第一报文发送给自身的控制面，并将所述第一报文添加 MC-LAG 标识后得到第二报文，将所述第二报文通过数据转发通道发送给所述虚拟集群中的第二网络设备的数据面，所述第二网络设备的数据面从所述第二报文中删除所述 MC-LAG 标识，得到所述第一报文后，确定所述第二网络上所述第一报文的入端口，并将确定的入端口以及所述第一报文发送给所述第二网络设备的控制面，所述第一网络设备的控制面根据所述第一网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，所述第二网络设备的控制面根据所述确定的入端口以及所述第一报文生成转发表项。本发明实施例通过虚拟集群中的多个网络设备之间的数据转发通道实现所述多个网络设备之间的转发表的一致性，不需要所述多个网络设备支持相同的软件版本，也不需要所述多个网络设备的软件版本互相兼容，能够避免所述多个网络设备没有同时升级软件版本时，导致的所述至少两个网络设备上的转发表不一致的问题，进而能够避免流量转发错误和用户业务中断。

附图说明

为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图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的前提下，还可以根据这些附图获得其他的附图。

图 1 为现有技术中 MC-LAG 的结构示意图；

图 2 为本发明实施例提供的转发表同步的系统的结构示意图；

图 3 为本发明实施例提供的一种转发表同步的方法的流程示意图；

图 4 为本发明实施例提供的另一种转发表同步的方法的流程示意图；

图 5 为本发明实施例提供的另一种转发表同步的方法的信令示意图；

图 6 为本发明实施例提供的一种网络设备的结构示意图；

图 7 为图 6 所示网络设备的硬件结构示意图；

图 8 为本发明实施例提供的另一种网络设备的结构示意图；

图 9 为图 8 所示网络设备的硬件结构示意图。

具体实施方式

为使本发明实施例的目的、技术方案和优点更加清楚，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作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。

如图2所示，为本发明实施例提供的转发表同步的系统的结构示意图。所述系统包括多归接入设备以及至少两个网络设备，每个网络设备均包括控制面和数据面。所述至少两个网络设备包括第一网络设备和第二网络设备。所述第一网络设备和第二网络设备组成虚拟集群，所述第一网络设备和所述第二网络设备具有相同的虚拟媒体访问控制MAC地址和相同的虚拟IP地址。所述第一网络设备为所述虚拟集群中的任意一个网络设备，所述第二网络设备为所述虚拟集群中除所述第一网络设备之外的网络设备。连接所述第一网络设备的第一端口与所述多归接入设备的第一链路，以及连接所述第二网络设备的第二端口与所述多归接入设备的第二链路组成一个MC-LAG，其MC-LAG标识为MC-LAG1，所述第一端口和第二端口为所述MC-LAG1的成员端口。所述多归接入设备通过MC-LAG1与所述虚拟集群通信。所述第一网络设备和第二网络设备上分别预先配置有MC-LAG表，所述MC-LAG表用于存储每个MC-LAG的MC-LAG标识与所述每个MC-LAG的成员端口的对应关系，在图2所示的实施例中，所述第一网络设备和第二网络设备上的MC-LAG表均包括表项{MC-LAG1，第一端口，第二端口}。所述第一网络设备通过数据转发通道与所述第二网络设备通信。

基于图2所示的系统，本发明实施例提供一种转发表同步的方法，应用于虚拟集群中，所述虚拟集群至少包括第一网络设备和第二网络设备，所述第一网络设备通过数据转发通道与所述第二网络设备通信，所述第一网络设备和第二网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组MC-LAG，如图3所示，所述方法包括：

301、虚拟集群中的第一网络设备的数据面接收多归接入设备发送的第一报文。

多归接入设备需要访问网络的时候，根据负载分担规则选择MC-LAG中的一条链路向虚拟集群发送第一报文。本实施例中，所述多归接入设备选择的链路为连接所述第一网络设备的链路，通过连接所述第一网络设备的链路，所述第一报文被从所述多归接入设备发送给所述第一网络设备的数据面。所述第一报文可以为广播报文，具体可以为ARP报文或者动态主机配置协议（英文：Dynamic Host Configuration Protocol，简称：

DHCP）报文。

302、所述第一网络设备的数据面将所述第一网络设备上所述第一报文的入端口以及所述第一报文发送给所述第一网络设备的控制面。

第一网络设备的数据面收到所述第一报文后，判断所述第一报文是否匹配控制面下发的报文选择规则，如果所述第一报文匹配所述控制面下发的报文选择规则，将所述第一网络设备上所述第一报文的入端口以及所述第一报文发送给所述第一网络设备的控制面。所述报文选择规则包括：选择特定类型的报文，例如，ARP报文或者DHCP报文。所述第一报文匹配所述控制面下发的报文选择规则，可以是，所述第一报文为ARP报文或者DHCP报文。所述报文选择规则还可以是：选择包括特定信息的报文，例如，选择IP地址、MAC地址或者VLAN ID属于某一范围或者为特定值的报文。所述第一报文匹配所述控制面下发的报文选择规则，可以是，所述第一报文的IP地址、MAC地址或者VLAN ID属于所述范围或者为所述特定值。将所述第一网络设备上所述第一报文的入端口以及所述第一报文发送给所述第一网络设备的控制面可以包括：将所述第一网络设备上所述第一报文的入端口以及所述第一报文分别发送给所述第一网络设备的控制面，或者将所述第一网络设备上所述第一报文的入端口添加到所述第一报文中，将所述添加了入端口的第一报文发送给所述第一网络设备的控制面。

303、所述第一网络设备的数据面为所述第一报文添加MC-LAG标识，得到第二报文，通过数据转发通道向所述虚拟集群中的第二网络设备的数据面发送所述第二报文，所述MC-LAG标识用于指示所述第一网络设备上所述第一报文的入端口为所述MC-LAG的成员端口。

具体地，所述第一网络设备根据所述第一网络设备上所述第一报文的入端口，即图2中的第一端口，以及预先配置的包括MC-LAG标识与端口的对应关系的MC-LAG表，获取所述入端口对应的MC-LAG标识，为所述第一报文添加所述获取的MC-LAG标识，得到所述第二报文。上述步骤302和步骤303并没有执行顺序的限制。

304、所述第一网络设备的控制面根据所述第一网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述第一网络设备的控制面中的第一转发表，将所述转发表项发送给所述第一网络设备的数据面。

根据所述第一网络设备上所述第一报文的入端口以及所述第一报文生成的转发表项为对应所述第一报文的转发表项。所述转发表项可以包括所述第一报文中包括的IP地址，MAC地址，虚拟局域网（英文：virtual local area network identifier，简称：

VLAN ID）中的至少一项。其中，所述转发表包括MAC地址表， ARP表，组播表，路由表中的至少一种。

305、所述第一网络设备的数据面接收所述第一网络设备的控制面发送的所述转发表项，根据所述转发表项更新所述第一网络设备的数据面中的第二转发表。

经过更新，所述第一网络设备中的第一转发表和所述第二转发表具有相同的内容。

相应于图3所示的方法，本发明实施例还提供另一种转发表同步的方法，应用于虚拟集群中，所述虚拟集群至少包括第一网络设备和第二网络设备，所述第一网络设备通过数据转发通道与所述第二网络设备通信，所述第一网络设备和第二网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组MC-LAG，如图4所示，所述方法包括：

401、虚拟集群中的第二网络设备的数据面通过数据转发通道接收所述虚拟集群中的第一网络设备的数据面发送的第二报文，所述第二报文是由所述第一网络设备为多归接入设备发送的第一报文添加MC-LAG标识后生成的，所述MC-LAG标识用于指示所述第一网络设备上所述第一报文的入端口为所述MC-LAG的成员端口。

402、所述第二网络设备的数据面根据所述第二报文，得到所述第一报文。

其中，所述第二网络设备的数据面根据所述第二报文，得到所述第一报文包括：

所述第二网络设备的数据面从所述第二报文中删除所述MC-LAG标识，得到所述第一报文。

403、所述第二网络设备的数据面确定所述第二网络设备上所述第一报文的入端口，并将所述确定的入端口和所述第一报文发送给所述第二网络设备的控制面。

所述确定所述第二网络设备上所述第一报文的入端口，包括：

所述第二网络设备的数据面根据所述第二报文获取所述MC-LAG标识，根据预先配置的包括MC-LAG标识与端口的对应关系的MC-LAG表，获取对应所述MC-LAG标识且属于所述第二网络设备的端口，将所述获取到的端口作为所述第二网络设备上所述第一报文的入端口。所述确定的入端口为所述第二网络设备上的所述MC-LAG的成员端口。

其中，将所述确定的入端口和所述第一报文发送给所述第二网络设备的控制面包括：将所述确定的入端口和所述第一报文分别发送给所述第二网络设备的控制面；或者，将所述确定的入端口添加到所述第一报文中，将所述添加了所述确定的入端口的第一报文发送给所述第二网络设备的控制面。

404、所述第二网络设备的控制面根据所述确定的入端口以及所述第一报文生成转

发表项，根据所述转发表项更新所述控制面中的第一转发表，并将所述转发表项发送给所述第二网络设备的数据面。

其中，所述第二网络设备的控制面生成转发表项的规则和所述第一网络设备的控制面生成转发表项的规则相同。因此，所述第二网络设备的控制面根据所述第一报文生成的转发表项除了入端口和出端口外，其他内容与所述第一网络设备的控制面根据所述第一报文生成的转发表项相同。此时，即认为所述第一网络设备的控制面和所述第二网络设备的控制面生成了一致的转发表项。

405、所述第二网络设备的数据面接收所述第二网络设备的控制面发送的所述转发表项，根据所述转发表项更新所述第二网络设备的数据面中的第二转发表。

经过更新，所述第二网络设备中的第一转发表和所述第二转发表具有相同的内容。

通过本发明图3和图4的实施例，第一网络设备和第二网络设备均生成了针对所述第一报文所属的业务流的转发表项。

如图5所示，本发明实施例还提供了另一种转发表同步的方法，所述方法包括：

501、第一网络设备的数据面通过第一端口接收多归接入设备发送的第一报文。

其中，所述第一报文可以数据报文或者协议报文。所述第一端口和第二网络设备上的第二端口为同一个MC-LAG的成员端口。所述第一网络设备和第二网络设备组成虚拟集群。

502、所述第一网络设备的数据面将所述第一端口以及所述第一报文发送给所述第一网络设备的控制面。

503、所述第一网络设备的数据面为所述第一报文添加MC-LAG标识，得到第二报文，并通过所述第一网络设备上的第三端口向第二网络设备的数据面发送所述第二报文。

504、所述第一网络设备的控制面根据根据所述第一网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述第一网络设备的控制面中的第一转发表。

505、所述第一网络设备的控制面将所述转发表项发送给所述第一网络设备的数据面。

506、所述第一网络设备的数据面接收所述第一网络设备的控制面发送的所述转发表项，根据所述转发表项更新所述第一网络设备的数据面中的第二转发表。

507、所述第二网络设备的数据面通过第四端口接收到所述第一网络设备的数据面发送的第二报文后，根据所述第二报文，得到所述第一报文。

508、所述第二网络设备的数据面确定所述第二网络设备上所述第一报文的入端口，并将所述确定的入端口和所述第一报文发送给所述第二网络设备的控制面。

509、所述第二网络设备的控制面根据所述确定的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述控制面中的第一转发表。

510、所述第二网络设备的控制面将所述转发表项发送给所述第二网络设备的数据面。

511、所述第二网络设备的数据面接收所述第二网络设备的控制面发送的所述转发表项，根据所述转发表项更新所述第二网络设备的数据面中的第二转发表。

上述各步骤中，涉及第一网络设备的步骤，可参考对图3的描述，涉及第二网络设备的步骤，可参考对图4的描述。其中，上述步骤502-506和上述步骤503、506-511之间，并没有执行上的先后顺序。

本发明上述实施例中，虚拟集群中的第一网络设备的数据面收到多归接入设备发送的第一报文后，将该第一报文发送给自身的控制面，并将所述第一报文添加 MC-LAG 标识后得到第二报文，将所述第二报文通过数据转发通道发送给所述虚拟集群中的第二网络设备的数据面，所述第二网络设备的数据面从所述第二报文中删除所述 MC-LAG 标识，得到所述第一报文后，确定所述第二网络上所述第一报文的入端口，并将确定的入端口以及所述第一报文发送给所述第二网络设备的控制面，所述第一网络设备的控制面根据所述第一网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，所述第二网络设备的控制面根据所述确定的入端口以及所述第一报文生成转发表项。本发明实施例通过虚拟集群中的多个网络设备之间的数据转发通道实现所述多个网络设备之间的转发表的一致性，不需要所述多个网络设备支持相同的软件版本，也不需要所述多个网络设备的软件版本互相兼容，能够避免所述多个网络设备没有同时升级软件版本时，导致的所述至少两个网络设备上的转发表不一致的问题，进而能够避免流量转发错误和用户业务中断。

基于图 2 所示的系统以及图 3 和图 5 所示的方法，本发明实施例进一步提供一种网络设备 600，所述网络设备 600 和另一网络设备组成虚拟集群，所述网络设备和另一网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组 MC-LAG。如图 6 所示，所述网络设备 600 包括控制面 601 和数据面 602。

所述数据面601用于接收所述多归接入设备发送的第一报文；将所述网络设备上所

述第一报文的入端口以及所述第一报文发送给所述控制面；为所述第一报文添加MC-LAG标识，得到第二报文，通过数据转发通道向所述另一网络设备的数据面发送所述第二报文，所述MC-LAG标识用于指示所述网络设备上所述第一报文的入端口为所述MC-LAG的成员端口。

所述控制面602用于根据所述网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述控制面中的第一转发表，将所述转发表项发送给所述数据面601。

所述数据面601还用于接收所述控制面发送的所述转发表项，根据所述转发表项更新所述数据面中的第二转发表。

其中，为所述第一报文添加 MC-LAG 标识时，所述数据面 601 具体用于根据所述网络设备上所述第一报文的入端口以及预先配置的包括 MC-LAG 标识与端口的对应关系的 MC-LAG 表，获取所述入端口对应的 MC-LAG 标识，为所述第一报文添加所述获取的 MC-LAG 标识。

其中，将所述网络设备上所述第一报文的入端口以及所述第一报文发送给所述控制面 602 时，所述数据面 601 具体用于判断所述第一报文是否匹配控制面 602 下发的报文选择规则，如果所述第一报文匹配所述报文选择规则，将所述网络设备上所述第一报文的入端口以及所述第一报文发送给所述控制面 602。

本发明实施例还提供了一种网络设备 700，如图 7 所示，所述装置 700 包括：

处理器(英文：processor)710，通信接口(英文：communications interface)720，存储器(英文：memory)730，总线 740。

处理器 710，通信接口 720，存储器 730 通过总线 740 进行相互间的通信。

通信接口 720，用于与外部网元通信。在一个实施方式中，所述通信接口 720 用于与多归接入设备、虚拟集群中的其他网络设备通信。通信接口 720 可以由光收发器，电收发器，无线收发器或其任意组合实现。例如，光收发器可以是小封装可插拔（英文：small form-factor pluggable transceiver，缩写：SFP）收发器（英文：transceiver），增强小封装可插拔（英文：enhanced small form-factor pluggable，缩写：SFP+）收发器或 10 吉比特小封装可插拔（英文：10 Gigabit small form-factor pluggable，缩写：XFP）收发器。电收发器可以是以太网（英文：Ethernet）网络接口控制器（英文：network interface controller，缩写：NIC)。无线收发器可以是无线网络接口控制器（英文：wireless network interface controller，缩写：WNIC)。通信接口 720

可以包括多个物理接口，例如通信接口 720 包括多个以太网接口。

处理器 710，用于执行程序 732。

具体地，程序 732 可以包括程序代码，所述程序代码包括计算机操作指令。

处理器 710 可能是中央处理器（英文：central processing unit，缩写：CPU），或者是专用集成电路（英文：application-specific integrated circuit，缩写：ASIC）。

存储器 730，用于存放程序 732。存储器 730 可以包括易失性存储器（英文：volatile memory），例如随机存取存储器（英文：random-access memory，缩写：RAM）；存储器 730 也可以包括非易失性存储器（英文：non-volatile memory），例如只读存储器（英文：read-only memory，缩写：ROM），快闪存储器（英文：flash memory），硬盘（英文：hard disk drive，缩写：HDD）或固态硬盘（英文：solid-state drive，缩写：SSD）；存储器 730 还可以包括上述种类的存储器的组合。

处理器 710 用于调用所述存储器中的程序 732，根据程序 732 执行图 3 以及图 5 中由第一网络设备所执行的步骤。

对图 6 以及图 7 描述不够详尽的地方，可以参考图 3 实施例中的描述。

基于图 2 所示的系统以及图 4 和图 5 所示的方法，本发明实施例进一步提供一种网络设备 800，所述网络设备 800 和另一网络设备组成虚拟集群，所述网络设备和另一网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组 MC-LAG。如图 8 所示，所述网络设备 800 包括控制面 801 和数据面 802。

所述数据面802用于通过所述数据转发通道接收所述另一网络设备的数据面发送的第二报文，所述第二报文是由所述另一网络设备为所述多归接入设备发送的第一报文添加MC-LAG标识后生成的，所述MC-LAG标识用于指示所述另一网络设备上所述第一报文的入端口为所述MC-LAG的成员端口；根据所述第二报文，得到所述第一报文；确定所述网络设备上所述第一报文的入端口，并将所述确定的入端口和所述第一报文发送给所述控制面801；

所述控制面801用于根据所述确定的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述控制面801中的第一转发表，并将所述转发表项发送给所述数据面802；

所述数据面802还用于接收所述控制面801发送的所述转发表项，根据所述转发表项更新所述数据面802中的第二转发表。

其中，在根据所述第二报文，得到所述第一报文时，所述数据面802具体用于从所述第二报文中删除所述MC-LAG标识，得到所述第一报文。

其中，在确定所述网络设备上所述第一报文的入端口时，所述数据面802具体用于：

根据所述第二报文获取所述 MC-LAG 标识，根据预先配置的包括 MC-LAG 标识与端口的对应关系的 MC-LAG 表，获取对应所述 MC-LAG 标识且属于所述网络设备的端口，将所述获取到的端口作为所述网络设备上所述第一报文的入端口。

本发明实施例还提供了一种网络设备 900，如图 9 所示，所述装置 900 包括：

处理器(英文：processor)910，通信接口(英文：communications interface)920，存储器(英文：memory)930，总线 940。

处理器 910，通信接口 920，存储器 930 通过总线 940 进行相互间的通信。

通信接口 920，用于与外部网元通信。在一个实施方式中，所述通信接口 920 用于与多归接入设备、虚拟集群中的其他网络设备通信。通信接口 920 的实现可以参考通信接口 620。

处理器 910，用于执行程序 932。

具体地，程序 932 可以包括程序代码，所述程序代码包括计算机操作指令。

处理器 910 可能是中央处理器（英文：central processing unit，缩写：CPU），或者是专用集成电路(英文：application-specific integrated circuit，缩写：ASIC)。

存储器 930，用于存放程序 932。存储器 930 的实现可以参考存储器 630。

处理器 910 用于调用所述存储器中的程序 932，根据程序 932 执行图 4 及图 5 中第二网络设备所执行的步骤。

对图 8 以及图 9 描述不够详尽的地方，可以参考图 4 及图 5 实施例中的描述。

本发明上述实施例中，虚拟集群中的第一网络设备的数据面收到多归接入设备发送的第一报文后，将该第一报文发送给自身的控制面，并将所述第一报文添加 MC-LAG 标识后得到第二报文，将所述第二报文通过数据转发通道发送给所述虚拟集群中的第二网络设备的数据面，所述第二网络设备的数据面从所述第二报文中删除所述 MC-LAG 标识，得到所述第一报文后，确定所述第二网络上所述第一报文的入端口，并将确定的入端口以及所述第一报文发送给所述第二网络设备的控制面，所述第一网络设备的控制面根据所述第一网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，所述第二网络设备的控制面根据所述确定的入端口以及所述第一报文生成转发表项。本发明实施例通过虚拟集群中的多个网络设备之间的数据转发通道实现所述多个网络设备之间的

转发表的一致性，不需要所述多个网络设备支持相同的软件版本，也不需要所述多个网络设备的软件版本互相兼容，能够避免所述多个网络设备没有同时升级软件版本时，导致的所述至少两个网络设备上的转发表不一致的问题，进而能够避免流量转发错误和用户业务中断。

以上所述仅为本发明的较佳实施例，并不用以限制本发明，凡在本发明的原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

权利要求

1、一种转发表同步的方法，应用于虚拟集群中，所述虚拟集群至少包括第一网络设备和第二网络设备，所述第一网络设备通过数据转发通道与所述第二网络设备通信，所述第一网络设备和第二网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组MC-LAG；

其特征在于，所述方法包括：

所述第一网络设备的数据面接收所述多归接入设备发送的第一报文；将所述第一网络设备上所述第一报文的入端口所述第一报文发送给所述第一网络设备的控制面；为所述第一报文添加MC-LAG标识，得到第二报文，通过所述数据转发通道向所述第二网络设备的数据面发送所述第二报文；所述MC-LAG标识用于指示所述第一网络设备上所述第一报文的入端口为所述MC-LAG的成员端口；

所述第一网络设备的控制面根据所述第一网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述第一网络设备的控制面中的第一转发表，将所述转发表项发送给所述第一网络设备的数据面；

所述第一网络设备的数据面接收所述第一网络设备的控制面发送的所述转发表项，根据所述转发表项更新所述第一网络设备的数据面中的第二转发表。

2、根据权利要求1所述的方法，其特征在于，所述第一网络设备的数据面为所述第一报文添加MC-LAG标识包括：

所述第一网络设备的数据面根据所述第一网络设备上所述第一报文的入端口以及预先配置的包括MC-LAG标识与端口的对应关系的MC-LAG表，获取所述入端口对应的MC-LAG标识，为所述第一报文添加所述获取的MC-LAG标识。

3、根据权利要求1或2所述的方法，其特征在于，所述第一网络设备的数据面将所述第一网络设备上所述第一报文的入端口以及所述第一报文发送给所述第一网络设备的控制面包括：

所述第一网络设备的数据面判断所述第一报文是否匹配所述第一网络设备的控制面下发的报文选择规则，如果所述第一报文匹配所述控制面下发的报文选择规则，将所述第一网络设备上所述第一报文的入端口以及所述第一报文发送给所述第一网络设备的控制面。

4、一种转发表同步的方法，应用于虚拟集群中，所述虚拟集群至少包括第一网络

设备和第二网络设备，所述第一网络设备通过数据转发通道与所述第二网络设备通信，所述第一网络设备和第二网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组MC-LAG；

其特征在于，所述方法包括：

所述第二网络设备的数据面通过所述数据转发通道接收所述第一网络设备的数据面发送的第二报文，所述第二报文是由所述第一网络设备为所述多归接入设备发送的第一报文添加MC-LAG标识后生成的，所述MC-LAG标识用于指示所述第一网络设备上所述第一报文的入端口为所述MC-LAG的成员端口；

所述第二网络设备的数据面根据所述第二报文，得到所述第一报文；

所述第二网络设备的数据面确定所述第二网络上所述第一报文的入端口，并将所述确定的入端口和所述第一报文发送给所述第二网络设备的控制面；

所述第二网络设备的控制面根据所述确定的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述控制面中的第一转发表，并将所述转发表项发送给所述第二网络设备的数据面；

所述第二网络设备的数据面接收所述第二网络设备的控制面发送的所述转发表项，根据所述转发表项更新所述第二网络设备的数据面中的第二转发表。

5、根据权利要求4所述的方法，其特征在于，所述根据所述第二报文，得到所述第一报文包括：

所述第二网络设备的数据面从所述第二报文中删除所述MC-LAG标识，得到所述第一报文。

6、根据权利要求4或5所述的方法，其特征在于，所述确定所述第二网络设备上所述第一报文的入端口包括：

所述第二网络设备的数据面根据所述第二报文获取所述MC-LAG标识，根据预先配置的包括MC-LAG标识与端口的对应关系的MC-LAG表，获取对应所述MC-LAG标识且属于所述第二网络设备的端口，将所述获取到的端口作为所述第二网络设备上所述第一报文的入端口。

7、一种网络设备，应用于虚拟集群中，所述虚拟集群至少包括所述网络设备和另一网络设备，所述网络设备通过数据转发通道与所述另一网络设备通信，所述网络设备和另一网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组MC-LAG；其特征在于，

所述网络设备包括控制面和数据面；

所述数据面用于接收所述多归接入设备发送的第一报文；将所述网络设备上所述第一报文的入端口以及所述第一报文发送给所述控制面；为所述第一报文添加MC-LAG标识，得到第二报文，通过所述数据转发通道向所述另一网络设备的数据面发送所述第二报文；所述MC-LAG标识用于指示所述网络设备上所述第一报文的入端口为所述MC-LAG的成员端口；

所述控制面用于根据所述网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述控制面中的第一转发表，将所述转发表项发送给所述数据面；

所述数据面还用于接收所述控制面发送的所述转发表项，根据所述转发表项更新所述数据面中的第二转发表。

8、根据权利要求7所述的网络设备，其特征在于，为所述第一报文添加MC-LAG标识时，所述数据面具体用于：

根据所述网络设备上所述第一报文的入端口以及预先配置的包括MC-LAG标识与端口的对应关系的MC-LAG表，获取所述入端口对应的MC-LAG标识，为所述第一报文添加所述获取的MC-LAG标识。

9、根据权利要求7或8所述的网络设备，其特征在于，将所述网络设备上所述第一报文的入端口以及所述第一报文发送给所述控制面时，所述数据面具体用于：

判断所述第一报文是否匹配所述控制面下发的报文选择规则，如果所述第一报文匹配所述控制面下发的报文选择规则，将所述网络设备上所述第一报文的入端口以及所述第一报文发送给所述控制面。

10、一种网络设备，应用于虚拟集群中，所述虚拟集群至少包括所述网络设备和另一网络设备，所述网络设备通过数据转发通道与所述另一网络设备通信，所述网络设备和另一网络设备分别通过第一链路和第二链路连接多归接入设备，所述第一链路和第二链路组成多机框链路聚合组MC-LAG；其特征在于，

所述网络设备包括控制面和数据面；

所述数据面通过所述数据转发通道接收所述另一网络设备的数据面发送的第二报文，所述第二报文是由所述另一网络设备为所述多归接入设备发送的第一报文添加MC-LAG标识后生成的，所述MC-LAG标识用于指示所述另一网络设备上所述第一报文的入端口为所述MC-LAG的成员端口；根据所述第二报文，得到所述第一报文；确定所述网络

设备上所述第一报文的入端口，并将所述确定的入端口和所述第一报文发送给所述控制面；

所述控制面用于根据所述确定的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述控制面中的第一转发表，并将所述转发表项发送给所述数据面；

所述数据面还用于接收所述控制面发送的所述转发表项，根据所述转发表项更新所述数据面中的第二转发表。

11、根据权利要求10所述的网络设备，其特征在于，在根据所述第二报文，得到所述第一报文时，所述数据面具体用于从所述第二报文中删除所述MC-LAG标识，得到所述第一报文。

12、根据权利要求10或11所述的网络设备，其特征在于，在确定所述网络设备上所述第一报文的入端口时，所述数据面具体用于：

根据所述第二报文获取所述MC-LAG标识，根据预先配置的包括MC-LAG标识与端口的对应关系的MC-LAG表，获取对应所述MC-LAG标识且属于所述网络设备的端口，将所述获取到的端口作为所述网络设备上所述第一报文的入端口。

13、一种转发表同步的系统，其特征在于，包括第一网络设备和第二网络设备；

所述第一网络设备为根据权利要求7-9中任意一项所述的网络设备；

所述第二网络设备为根据权利要求10-12中任意一项所述的网络设备。

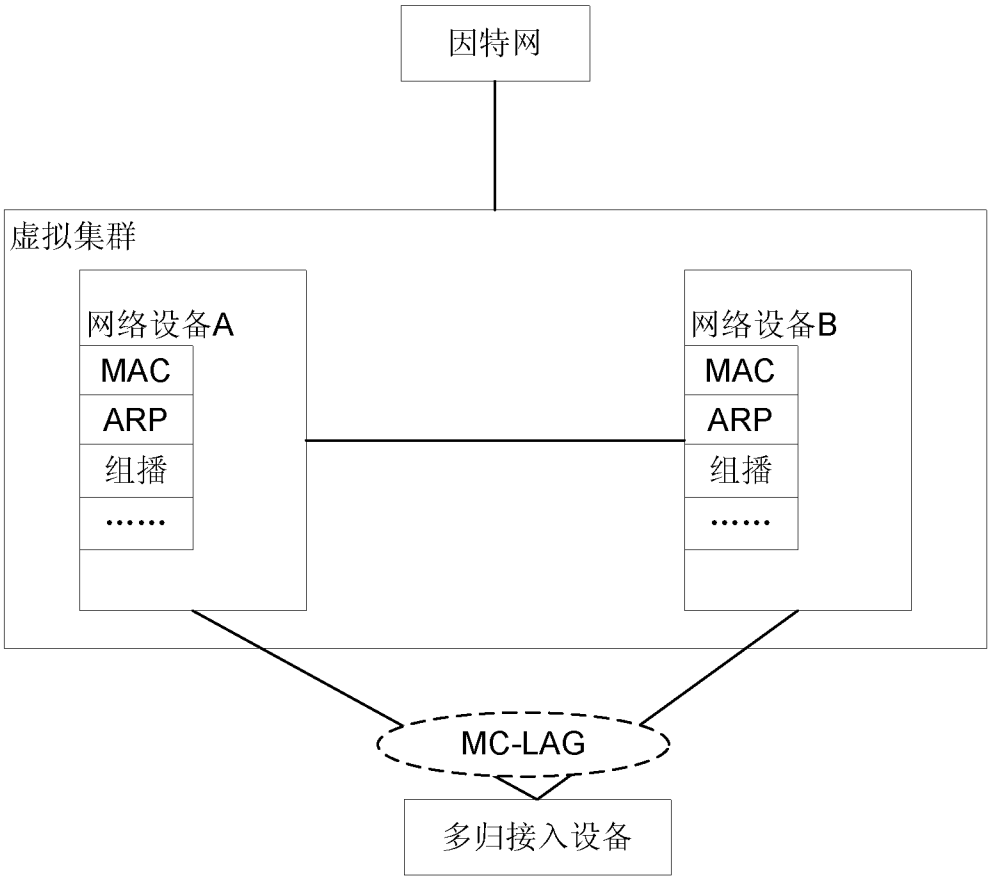

因特网
虚拟集群
网络设备A
MAC
ARP
组播
……
网络设备B
MAC
ARP
组播
……
MC-LAG
多归接入设备

图 1

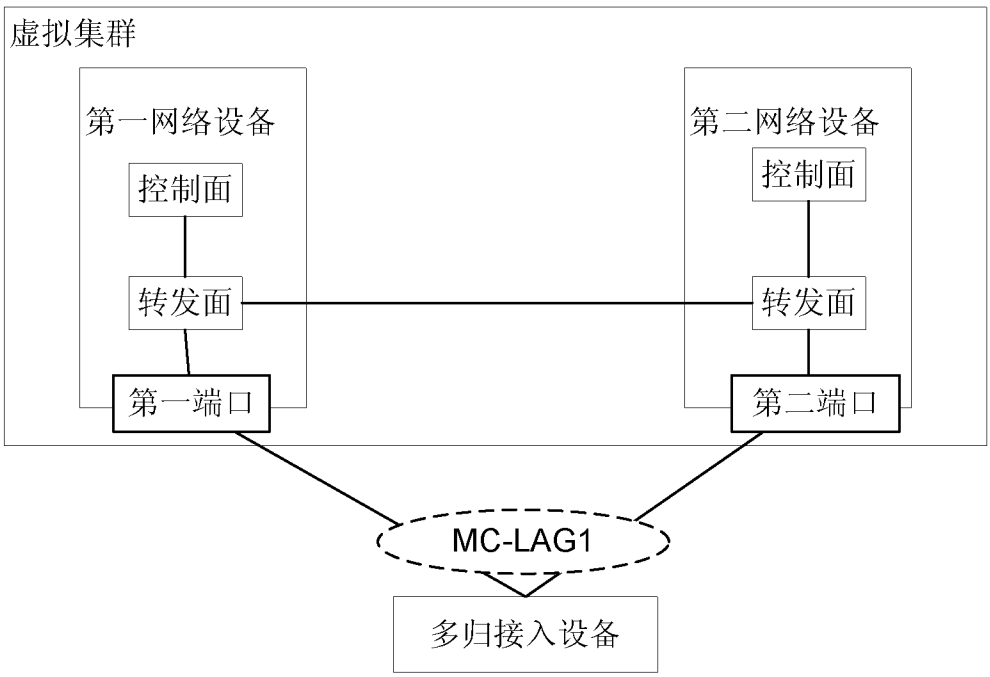

虚拟集群
第一网络设备
控制面
转发面
第一端口
第二网络设备
控制面
转发面
第二端口
MC-LAG1
多归接入设备

图 2

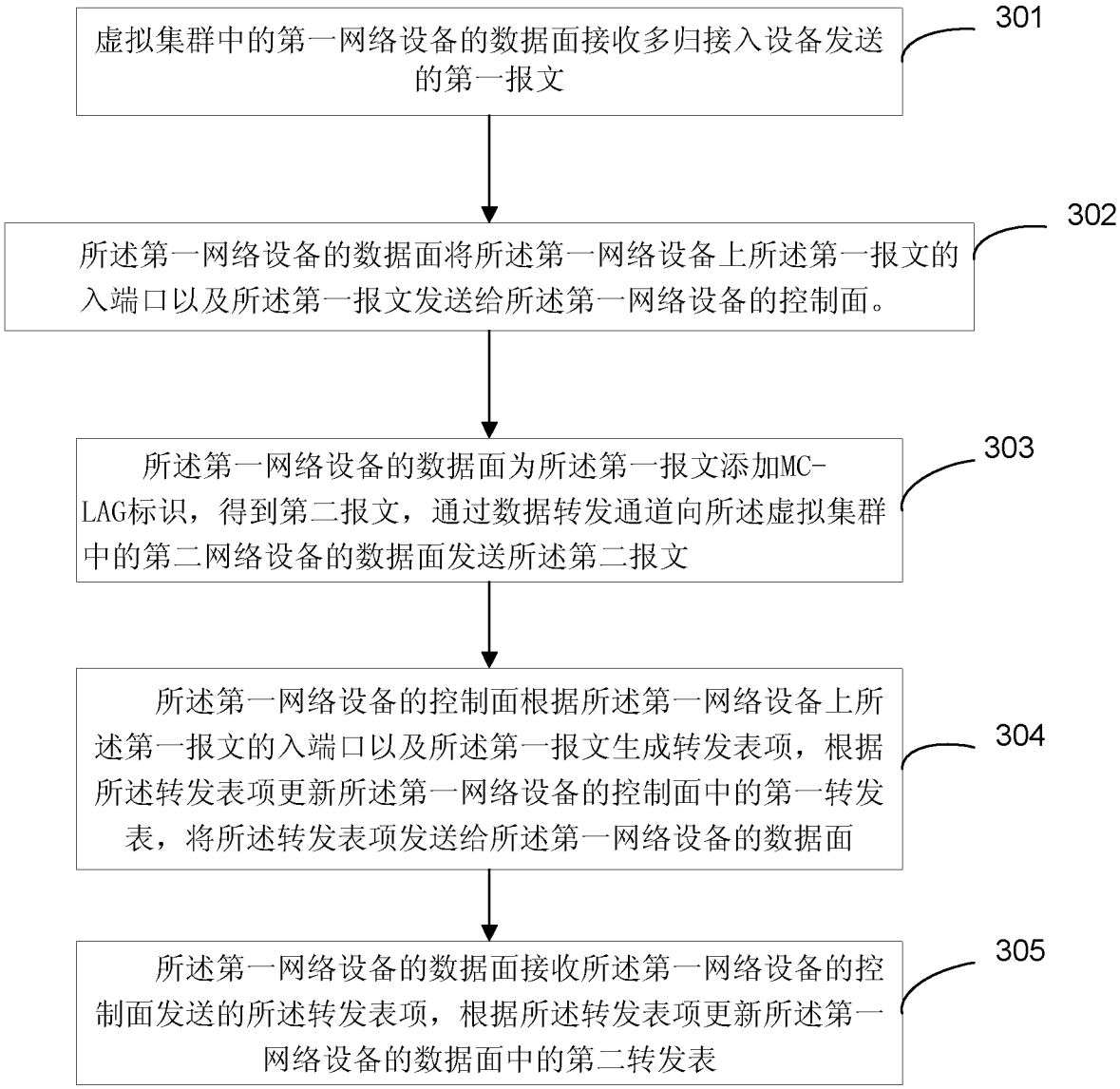

虚拟集群中的第一网络设备的数据面接收多归接入设备发送的第一报文
301
所述第一网络设备的数据面将所述第一网络设备上所述第一报文的入端口以及所述第一报文发送给所述第一网络设备的控制面。
302
所述第一网络设备的数据面为所述第一报文添加MC-LAG标识，得到第二报文，通过数据转发通道向所述虚拟集群中的第二网络设备的数据面发送所述第二报文
303
所述第一网络设备的控制面根据所述第一网络设备上所述第一报文的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述第一网络设备的控制面中的第一转发表，将所述转发表项发送给所述第一网络设备的数据面
304
所述第一网络设备的数据面接收所述第一网络设备的控制面发送的所述转发表项，根据所述转发表项更新所述第一网络设备的数据面中的第二转发表
305

图 3

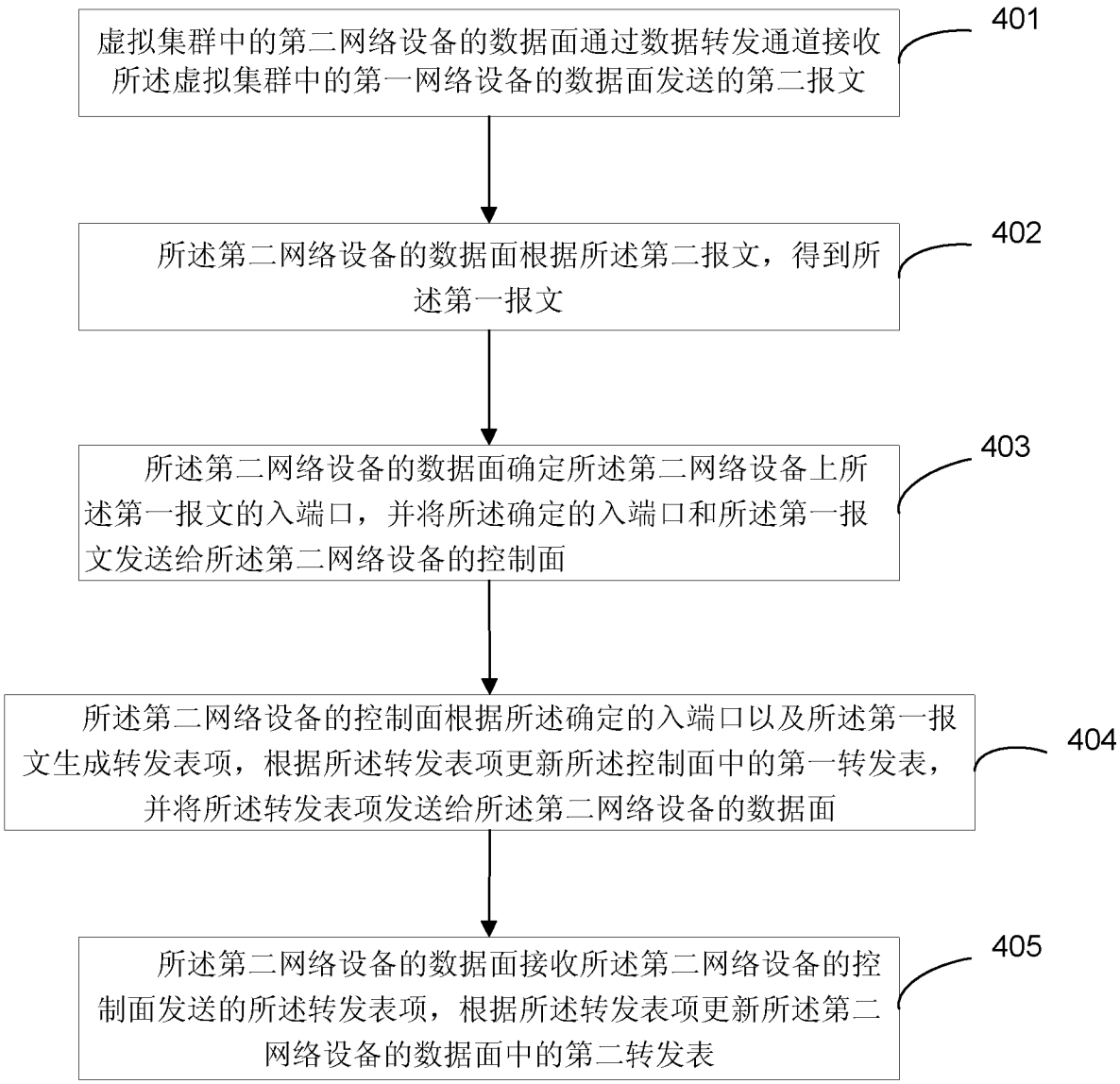

虚拟集群中的第二网络设备的数据面通过数据转发通道接收所述虚拟集群中的第一网络设备的数据面发送的第二报文
401
所述第二网络设备的数据面根据所述第二报文，得到所述第一报文
402
所述第二网络设备的数据面确定所述第二网络设备上所述第一报文的入端口，并将所述确定的入端口和所述第一报文发送给所述第二网络设备的控制面
403
所述第二网络设备的控制面根据所述确定的入端口以及所述第一报文生成转发表项，根据所述转发表项更新所述控制面中的第一转发表，并将所述转发表项发送给所述第二网络设备的数据面
404
所述第二网络设备的数据面接收所述第二网络设备的控制面发送的所述转发表项，根据所述转发表项更新所述第二网络设备的数据面中的第二转发表
405

图 4

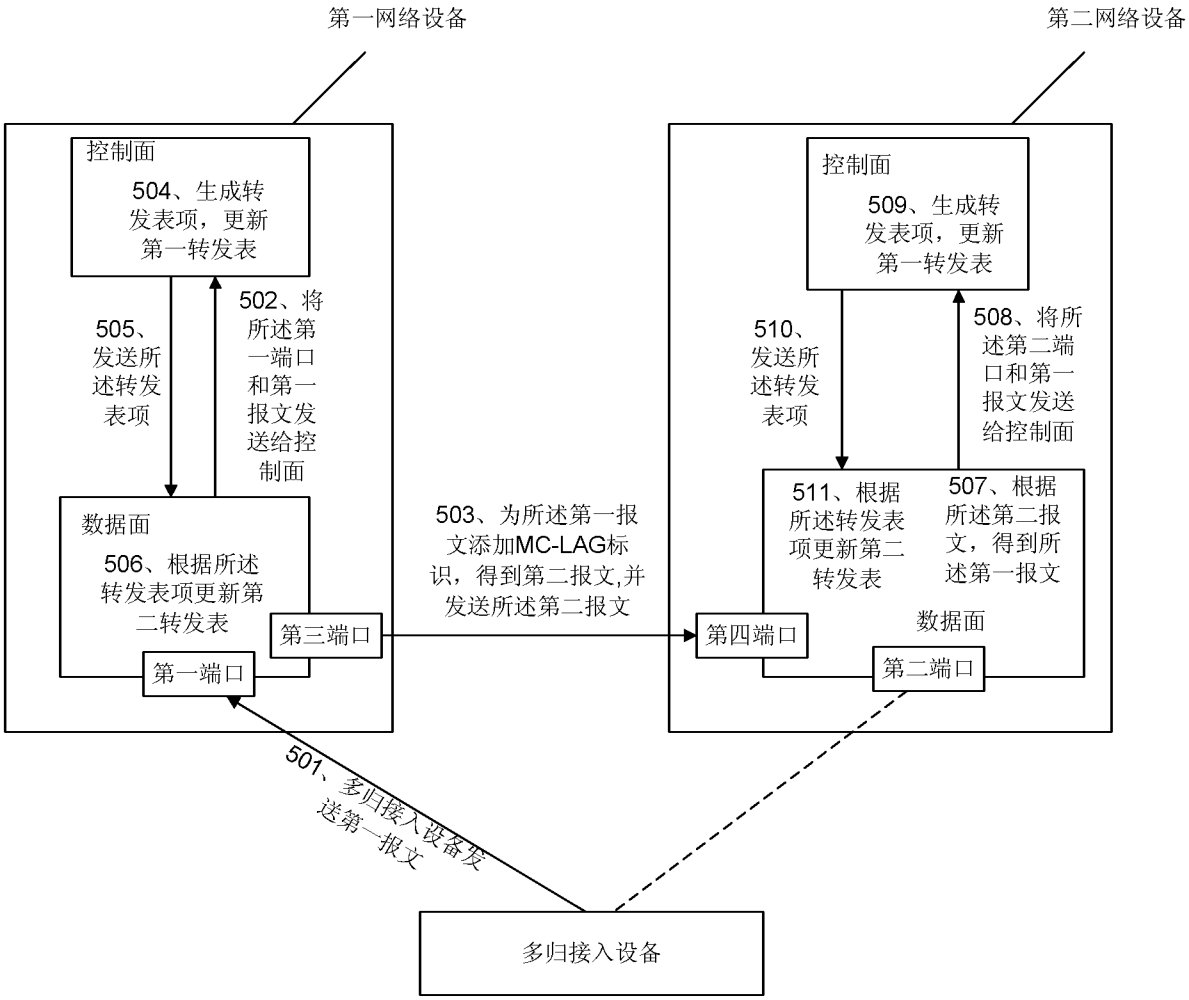

第一网络设备
第二网络设备
控制面
504、生成转发表项，更新第一转发表
505、发送所述转发表项
502、将所述第一端口和第一报文发送给控制面
数据面
506、根据所述转发表项更新第二转发表
第三端口
第一端口
503、为所述第一报文添加MC-LAG标识，得到第二报文,并发送所述第二报文
控制面
509、生成转发表项，更新第一转发表
510、发送所述转发表项
508、将所述第二端口和第一报文发送给控制面
511、根据所述转发表项更新第二转发表
507、根据所述第二报文，得到所述第一报文
数据面
第四端口
第二端口
501、多归接入设备发送第一报文
多归接入设备

图 5

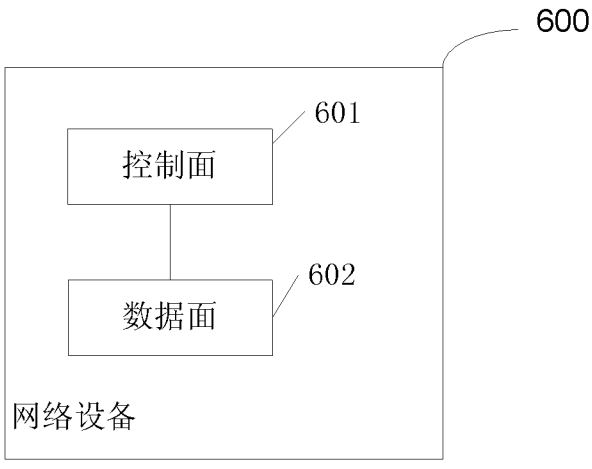

600
601
控制面
602
数据面
网络设备

图 6

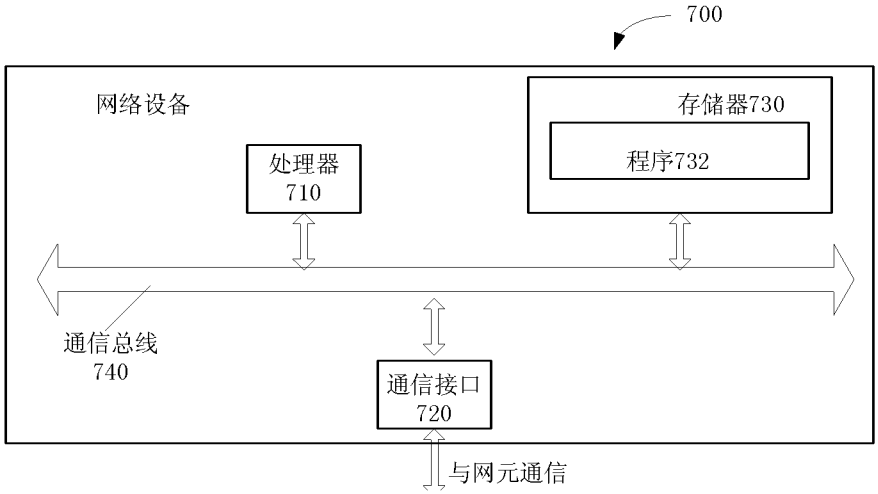

700
网络设备
存储器730
程序732
处理器
710
通信总线
740
通信接口
720
与网元通信

图 7

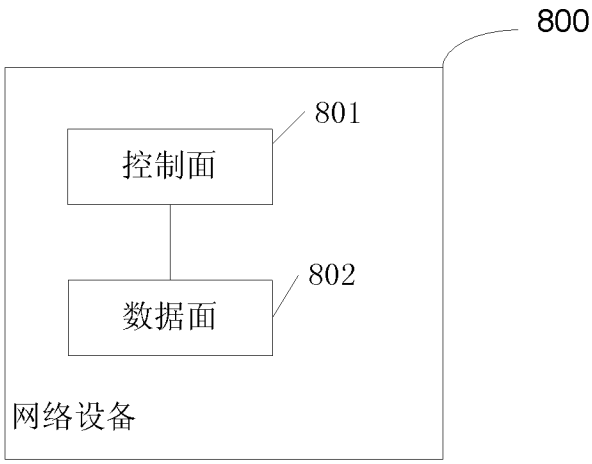

800
801
控制面
802
数据面
网络设备

图 8

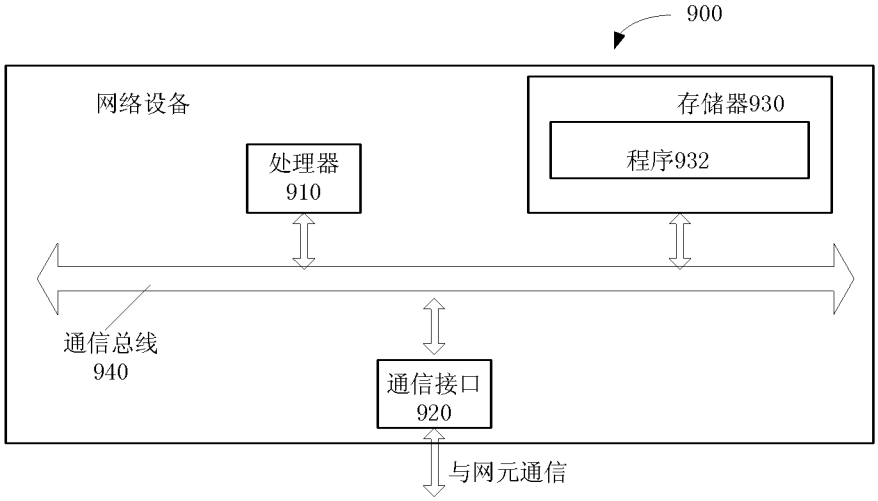

900
网络设备
存储器930
程序932
处理器
910
通信总线
940
通信接口
920
与网元通信

图 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2014/085944

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/757 (2013.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L 12/-; H04L 29/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

USTXT; CNTXT; CNABS; VEN; GOOGLE SCHOLAR; IEEE: LAG, ALG, cluster, transferring service flow, multi-normalization access device, multi-normalization, data layer, control layer, MAC, arp, multicast, route, mc 1w lag, update; group, link 1w aggregation 1w group, multi 1w chassis, port? 1w aggregation, aggregation 1w port?, data 1w plane, control 1w plane, transfer+ 1w table, lag 1w table, route 2w table, synchroniz+, message, multi chassis

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 103098424 A (ALCATEL-LUCENT), 08 May 2013 (08.05.2013), the whole document	1-13
A	CN 102148677 A (H3C TECHNOLOGIES CO., LIMITED), 10 August 2011 (10.08.2011), the whole document	1-13
A	CN 103414801 A (H3C TECHNOLOGIES CO., LIMITED), 27 November 2013 (27.11.2013), the whole document	1-13
A	CN 101877667 A (HUAWEI TECHNOLOGIES CO., LTD.), 03 November 2010 (03.11.2010), the whole document	1-13
A	CN 1725761 A (HANGZHOU HUAWEI 3COM TECHNOLOGY CO., LTD.), 25 January 2006 (25.01.2006), the whole document	1-13
A	CN 103166874 A (H3C TECHNOLOGIES CO., LIMITED), 19 June 2013 (19.06.2013), the whole document	1-13

☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 October 2014 (31.10.2014)	**06 February 2015 (06.02.2015)**
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer **GAO, Shengkai** Telephone No.: (86-10) **0512-88996426**

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2014/085944

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 101277219 A (NEC CORPORATION), 01 October 2008 (01.10.2008), the whole document	1-13
A	CN 103929461 A (H3C TECHNOLOGIES CO., LIMITED), 16 July 2014 (16.07.2014), the whole document	1-13
A	US 7903554 B1 (FORCE 10 NETWORKS INC.), 08 March 2011 (08.03.2011), the whole document	1-13
A	US 2011194562 A1 (FUJITSU NETWORK COMMUNICATIONS INC.), 11 August 2011 (11.08.2011), the whole document	1-13
A	US 2012236859 A1 (FORCE 10 NETWORKS INC.), 20 September 2012 (20.09.2012), the whole document	1-13
A	US 2012128005 A1 (HITACHI CABLE LTD.), 24 May 2012 (24.05.2012), the whole document	1-13
A	US 2014219081 A1 (DELL PRODUCTS LP), 07 August 2014 (07.08.2014), the whole document	1-13
T	HENK, B. et al. "Supporting Security Automation for Multi-Chassis Link Aggregation Groups via the Interconnected-asset Ontology" AVAILABILITY, RELIABILITY AND SECURITY (ARES), 2014 NINTH INTERNATIONAL CONFERENCE ON, 12 September 2014 (12.09.2014), pages 126-133	1-13

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2014/085944

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103098424 A	08 May 2013	WO 2012018521 A1	09 February 2012
		KR 20130032384 A	01 April 2013
		US 8488608 B2	16 July 2013
		US 8767735 B2	01 July 2014
		US 2012033541 A1	09 February 2012
		US 2012033672 A1	09 February 2012
		US 2012033669 A1	09 February 2012
		US 8462774 B2	11 June 2013
		JP 2013535922 A	12 September 2013
		EP 2601763 A1	12 June 2013
		US 2012033665 A1	09 February 2012
		US 2012033668 A1	09 February 2012
		US 2012033678 A1	09 February 2012
		US 2012182866 A1	19 July 2012
		US 2013064137 A1	14 March 2013
		US 2013073711 A1	21 March 2013
		US 2013077621 A1	28 March 2013
		US 8472447 B2	25 June 2013
		US 8582423 B2	12 November 2013
		WO 2012099946 A1	26 July 2012
		WO 2012099858 A1	26 July 2012
		KR 1455013 B1	27 October 2014
CN 102148677 A	10 August 2011	CN 102148677 B	26 June 2013
CN 103414801 A	27 November 2013	None	
CN 101877667 A	03 November 2010	CN 101877667 B	09 January 2013
		US 2010278180 A1	04 November 2010
		US 8289967 B2	16 October 2012
CN 1725761 A	25 January 2006	CN 100479465 C	15 April 2009
CN 103166874 A	19 June 2013	WO 2014154123 A1	02 October 2014
CN 101277219 A	01 October 2008	JP 4830942 B2	07 December 2011
		US 7835376 B2	16 November 2010
		JP 2008244907 A	09 October 2008
		CN 101277219 B	14 August 2013
		US 2008240133 A1	02 October 2008
CN 103929461 A	16 July 2014	None	

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2014/085944

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 7903554 B1	08 March 2011	None	
US 2011194562 A1	11 August 2011	US 8462636 B2	11 June 2013
US 2012236859 A1	20 September 2012	US 8649379 B2	11 February 2014
US 2012128005 A1	24 May 2012	US 8737412 B2	27 May 2014
		JP 2012114644 A	14 June 2012
		JP 5522003 B2	18 June 2014
		JP 2013070297 A	18 April 2013
US 2014219081 A1	07 August 2014	WO 2014120282 A1	07 August 2014

国际检索报告

国际申请号
PCT/CN2014/085944

A. 主题的分类

H04L 12/757(2013.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

H04L12/-; H04L29/-

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用）)

USTXT;CNTXT;CNABS;VEN;GOOGLE SCHOLAR;IEEE:LAG, 链路聚合组, ALG,集群,转发表, 转发业务流, 端口聚合, 聚合端口, 多归接入设备, 多归, 数据面, 数据层, 控制面, 控制层, MAC, arp, 组播, 路由, mc lw lag, 同步, 更新, lag表; group, link lw aggregation lw group, multi lw chassis, port? lw aggregation, aggregation lw port?, data lw plane, control lw plane, transfer+ lw table, lag lw table, route 2w table, synchroniz+, message,multi chassis

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 103098424 A (阿尔卡特朗讯公司) 2013年 5月 08日 (2013-05-08) 全文	1-13
A	CN 102148677 A (杭州华三通信技术有限公司) 2011年 8月 10日 (2011-08-10) 全文	1-13
A	CN 103414801 A (杭州华三通信技术有限公司) 2013年 11月 27日 (2013-11-27) 全文	1-13
A	CN 101877667 A (华为技术有限公司) 2010年 11月 03日 (2010-11-03) 全文	1-13
A	CN 1725761 A (杭州华为三康技术有限公司) 2006年 1月 25日 (2006-01-25) 全文	1-13
A	CN 103166874 A (杭州华三通信技术有限公司) 2013年 6月 19日 (2013-06-19) 全文	1-13

☑ 其余文件在C栏的续页中列出。　　☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2014年 10月 31日	2015年 2月 06日

ISA/CN的名称和邮寄地址	受权官员
中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国 传真号 (86-10)62019451	高胜凯 电话号码 (86-10)0512-88996426

表 PCT/ISA/210 (第2页) (2009年7月)

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 101277219 A (日本电气株式会社) 2008年 10月 01日 (2008-10-01) 全文	1-13
A	CN 103929461 A (杭州华三通信技术有限公司) 2014年 7月 16日 (2014-07-16) 全文	1-13
A	US 7903554 B1 (FORCE 10 NETWORKS INC) 2011年 3月 08日 (2011-03-08) 全文	1-13
A	US 2011194562 A1 (FUJITSU NETWORK COMMUNICATIONS INC) 2011年 8月 11日 (2011-08-11) 全文	1-13
A	US 2012236859 A1 (FORCE 10 NETWORKS INC) 2012年 9月 20日 (2012-09-20) 全文	1-13
A	US 2012128005 A1 (HITACHI CABLE LTD) 2012年 5月 24日 (2012-05-24) 全文	1-13
A	US 2014219081 A1 (DELL PRODUCTS LP) 2014年 8月 07日 (2014-08-07) 全文	1-13
T	Henk Birkholz等. ""Supporting Security Automation for Multi-Chassis Link Aggregation Groups via the Interconnected-asset Ontology"" Availability, Reliability and Security (ARES), 2014 Ninth International Conference on, 2014年 9月 12日 (2014-09-12), 126 - 133页	1-13

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2014/085944

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	103098424	A	2013年 5月 08日	WO	2012018521	A1	2012年 2月 09日
				KR	20130032384	A	2013年 4月 01日
				US	8488608	B2	2013年 7月 16日
				US	8767735	B2	2014年 7月 01日
				US	2012033541	A1	2012年 2月 09日
				US	2012033672	A1	2012年 2月 09日
				US	2012033669	A1	2012年 2月 09日
				US	8462774	B2	2013年 6月 11日
				JP	2013535922	A	2013年 9月 12日
				EP	2601763	A1	2013年 6月 12日
				US	2012033665	A1	2012年 2月 09日
				US	2012033668	A1	2012年 2月 09日
				US	2012033678	A1	2012年 2月 09日
				US	2012182866	A1	2012年 7月 19日
				US	2013064137	A1	2013年 3月 14日
				US	2013073711	A1	2013年 3月 21日
				US	2013077621	A1	2013年 3月 28日
				US	8472447	B2	2013年 6月 25日
				US	8582423	B2	2013年 11月 12日
				WO	2012099946	A1	2012年 7月 26日
				WO	2012099858	A1	2012年 7月 26日
				KR	1455013	B1	2014年 10月 27日
CN	102148677	A	2011年 8月 10日	CN	102148677	B	2013年 6月 26日
CN	103414801	A	2013年 11月 27日		无		
CN	101877667	A	2010年 11月 03日	CN	101877667	B	2013年 1月 09日
				US	2010278180	A1	2010年 11月 04日
				US	8289967	B2	2012年 10月 16日
CN	1725761	A	2006年 1月 25日	CN	100479465	C	2009年 4月 15日
CN	103166874	A	2013年 6月 19日	WO	2014154123	A1	2014年 10月 02日
CN	101277219	A	2008年 10月 01日	JP	4830942	B2	2011年 12月 07日
				US	7835376	B2	2010年 11月 16日
				JP	2008244907	A	2008年 10月 09日
				CN	101277219	B	2013年 8月 14日
				US	2008240133	A1	2008年 10月 02日
CN	103929461	A	2014年 7月 16日		无		
US	7903554	B1	2011年 3月 08日		无		
US	2011194562	A1	2011年 8月 11日	US	8462636	B2	2013年 6月 11日
US	2012236859	A1	2012年 9月 20日	US	8649379	B2	2014年 2月 11日
US	2012128005	A1	2012年 5月 24日	US	8737412	B2	2014年 5月 27日
				JP	2012114644	A	2012年 6月 14日
				JP	5522003	B2	2014年 6月 18日
				JP	2013070297	A	2013年 4月 18日
US	2014219081	A1	2014年 8月 07日	WO	2014120282	A1	2014年 8月 07日