

(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
2015年12月3日 (03.12.2015)



(10) 国际公布号
WO 2015/180153 A1

- (51) 国际专利分类号:
H04L 12/741 (2013.01)
- (21) 国际申请号: PCT/CN2014/078972
- (22) 国际申请日: 2014年5月30日 (30.05.2014)
- (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: 曾毓珑 (ZENG, Yulong); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 文白林 (WEN, Bailin); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 李建飞 (LI, Jianfei); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 北京同立钧成知识产权代理有限公司 (LEADER PATENT & TRADEMARK FIRM); 中国北京市海淀区西直门北大街32号枫蓝国际A座8F-6, Beijing 100082 (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, LT, 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, 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: CONSTRUCTION METHOD, DEVICE AND SYSTEM FOR MULTI-PATH FORWARDING RULES

(54) 发明名称: 一种多路径转发规则的构造方法、装置及系统

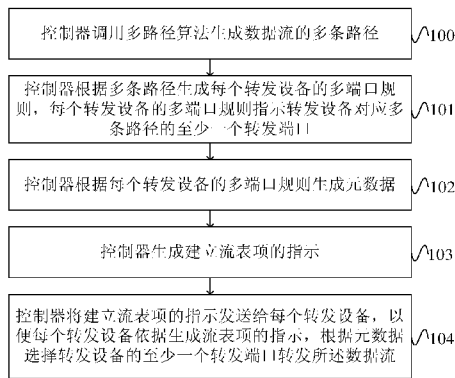


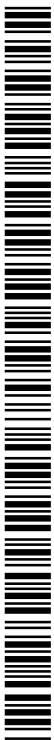
图2 / Fig. 2

- 100 A controller calls a multi-path algorithm to generate a plurality of paths of a data stream
- 101 The controller generates multi-port rules of each forwarding device according to the plurality of paths, wherein the multi-port rules of each forwarding device indicates at least one forwarding port of the forwarding device corresponding to the plurality of paths
- 102 The controller generates metadata according to the multi-port rules of each forwarding device
- 103 The controller generates an indication to generate a flow entry
- 104 The controller sends the indication to generate a flow entry to each forwarding device, so that each forwarding device selects at least one forwarding port of the forwarding device to forward the data stream according to the indication to generate a flow entry and according to the metadata

(57) Abstract: Provided are a construction method, device and system for multi-path forwarding rules. A controller calls a multi-path algorithm to generate a plurality of paths of a data stream; the controller generates multi-port rules of each forwarding device according to the plurality of paths; the controller generates metadata according to the multi-port rules of each forwarding device, wherein the metadata contains information for indicating at least one forwarding port of the forwarding device; the controller generates an indication to establish a flow entry; the controller sends the indication to generate a flow entry to each forwarding device, so that each forwarding device selects at least one forwarding port of the forwarding device to forward the data stream according to the indication to generate a flow entry and according to the metadata. Since the forwarding device can implement multiplexing on a port combination according to metadata, after generating a flow entry, the controller need not query a group that has been used and directly sends the flow entry to each forwarding device, so as to accelerate the decision time of the controller.

(57) 摘要:

[见续页]



WO 2015/180153 A1

本发明提供一种多路径转发规则的构造方法、装置及系统，控制器调用多路径算法生成数据流的多条路径；控制器根据多条路径生成每个转发设备的多端口规则；控制器根据每个转发设备的多端口规则生成元数据，元数据中包含指示转发设备的至少一个转发端口的信息；控制器生成建立流表项的指示；控制器将建立流表项的指示发送给每个转发设备，以便每个转发设备依据生成流表项的指示，根据元数据选择转发设备的至少一个转发端口转发数据流；由于转发设备根据元数据既可以对端口组合实现复用，因此，控制器在生成流表项后，不需要查询已经使用过的组，直接将流表项发送给每个转发设备，从而加快了控制器的决策时间。

一种多路径转发规则的构造方法、装置及系统

技术领域

- 5 本发明涉及网络通信技术领域，尤其涉及多路径转发规则的构造方法、装置及系统。

背景技术

- 随着网络技术的发展，软件定义网络（Software Defined Network, SDN）技术的提出是对网络架构的一次创新，其主要技术特征包括：转发控制分离、网络集中管控、支持开放编程以开放网络资源和能力。SDN的目标在于通过对网络架构的重构来简化网络的管控，加快新业务的部署。OpenFlow是一种主流的实现SDN的关键技术，其定义了网络控制器和网络转发设备间通信的接口和协议。基于OpenFlow的SDN网络中，引入组(group)的概念使一组端口集合作为一个单一实体用于包的转发，并使用组表（group table）来存储组表项，组表项中使用一个标识符ID唯一标识一个组表项，组类型（group type）代表该组支持的业务类型，计数器（Counters）用于针对组的流量统计，动作桶（action buckets）用于指定该组表项关联的操作桶（bucket）。不同的组类型可以支持不同的业务。一个操作桶项中包含桶的标识、权重和操作（如转发到某个端口），其中权重用于确定数据在各条转发路径上转发的比例。
- 10
15
20

- 在SDN网络中，构造一条多路径转发规则的通常方法是，1) 定义一个组，指定组类型为选择（Select），代表数据包转发将从多个操作桶中根据一定的规则/算法选择一个端口；指定其操作桶，例如指定多个转发端口号和权重。2) 构造流表项（flow entry），包含匹配域（Match Field）和操作域（Action），其中操作域部分定义为指向组表中的某个表项。
- 25

- 但是，转发设备的存储资源是有限的，组表支持的硬件流表条数也是有限的，因而组的复用对节省存储空间来说尤为重要。现有技术中，要实现组的复用，控制器需要查询已经使用过的组，对比历史组的操作桶中的端口组合是否和将建立的组一致，在一致的情况下使用历史组的标识作为新建组的标识使用。当使用多路径转发的组较多时，端口组合的情况种类较多，查询历史组的标识将产生不小的开销，影响到转发设备上流表建立的时间，降低
- 30

了控制器与转发设备的工作效率。

发明内容

5 本发明提供一种多路径转发规则的构造方法、装置及系统，用于提高多路径转发规则的构造效率。

本发明的第一个方面是提供一种多路径转发规则构造方法，包括：

控制器调用多路径算法生成数据流的多条路径；

10 所述控制器根据所述多条路径生成每个转发设备的多端口规则，每个所述转发设备的多端口规则指示所述转发设备对应所述多条路径的至少一个转发端口；

所述控制器根据每个所述转发设备的多端口规则生成元数据，所述元数据中包含指示所述转发设备的所述至少一个转发端口的信息；

15 所述控制器生成建立流表项的指示，所述建立流表项的指示中包含写元数据指示、跳转到多路径表指示和所述写元数据指示对应的所述元数据；

所述控制器将所述建立流表项的指示发送给每个所述转发设备，以便每个所述转发设备依据所述生成流表项的指示，根据所述元数据选择所述转发设备的所述至少一个转发端口转发所述数据流。

20 结合第一个方面，在第一种可能的实现方式中，所述元数据中包含端口信息标识区，所述端口信息标识区用于指示所述转发设备的所述至少一个转发端口的信息，所述端口信息标识区中的每个比特位对应指示一个所述转发端口。

25 结合第一个方面或第一个方面的第一种可能的实现方式，在第二种可能的实现方式中，所述控制器根据每个所述转发设备的多端口规则生成元数据之前，还包括：

所述控制器判断每个所述转发设备的多端口规则对应的分流比例是否为等比，若是等比，则执行所述控制器根据每个所述转发设备的多端口规则生成所述元数据的步骤。

30 本发明的第二个方面是提供一种多路径转发规则构造方法，包括：转发设备接收控制器发送的建立流表项的指示，所述建立流表项的指

示中包含写元数据指示、跳转到多路径表指示和所述写元数据指示对应的所述元数据；

所述转发设备根据所述建立流表项的指示建立第一张表的流表项，所述第一张表的流表项中的动作域为写入所述写元数据指示对应的所述元数据、以及跳转到多路径表，其中，所述第一张流表指的是所述转发设备中的入口流表；

所述转发设备根据所述写元数据指示建立所述多路径表的流表项，所述多路径表的流表项中，匹配域为所述元数据，动作域为从所述元数据对应的至少一个转发端口发送数据流。

结合第二个方面，在第一种可能的实现方式中，所述元数据中包含端口信息标识区，所述端口信息标识区用于指示所述转发设备的所述至少一个转发端口的信息，所述端口信息标识区中的每个比特位对应指示一个所述转发端口；

所述动作域中为根据所述端口信息标识区中指示的所述转发设备的所述至少一个转发端口的信息一一对应指示的所述转发端口作为多路径转发的端口，发送所述数据流。

结合第二个方面的第一种可能的实现方式，在第二种可能的实现方式中，所述转发设备根据所述写元数据指示建立所述多路径表的流表项，包括；

所述转发设备查询所述多路径表中是否存在包含所述元数据的流表项；

若不存在，则所述转发设备根据所述写元数据指示，提取所述元数据中所述端口信息标识区指示的所述转发设备的所述至少一个转发端口的信息，对应生成所述多路径表的流表项中的所述动作域。

本发明的第三个方面是提供一种控制器，包括：

处理模块，用于调用多路径算法生成数据流的多条路径，根据所述多条路径生成每个转发设备的多端口规则，每个所述转发设备的多端口规则指示所述转发设备对应所述多条路径的至少一个转发端口，根据每个所述转发设备的多端口规则生成元数据，所述元数据中包含指示所述转发设备的所述至少一个转发端口的信息，生成建立流表项的指示，所述建立流表

项的指示中包含写元数据指示、跳转到多路径表指示和所述写元数据指示对应的所述元数据；

5 发送模块，用于将所述建立流表项的指示发送给每个所述转发设备，以便每个所述转发设备依据所述建立流表项的指示，根据所述元数据选择所述转发设备的所述至少一个转发端口转发所述数据流。

结合第三个方面，在第一种可能的实现方式中，所述元数据中包含端口信息标识区，所述端口信息标识区用于指示所述转发设备的所述至少一个转发端口的信息，所述端口信息标识区中的每个比特位对应指示一个所述转发端口。

10 结合第三个方面或第三个方面的第一种可能的实现方式，在第二种可能的实现方式中，所述处理模块，还用于在根据每个所述转发设备的多端口规则生成元数据之前，判断每个所述转发设备的多端口规则对应的分流比例是否为等比，若是等比，则根据每个所述转发设备的多端口规则生成所述元数据。

15 本发明的第四个方面是提供一种转发设备，包括：

接收模块，用于接收控制器发送的建立流表项的指示，所述建立流表项的指示中包含写元数据指示、跳转到多路径表指示和所述写元数据指示对应的所述元数据；

20 处理模块，用于根据所述建立流表项的指示建立第一张表的流表项，所述第一张表的流表项中的动作域为写入所述写元数据指示对应的所述元数据、以及跳转到多路径表；还用于根据所述写元数据指示建立所述多路径表的流表项，所述多路径表的流表项中，匹配域为所述元数据，动作域为从所述元数据对应的至少一个转发端口发送数据流，所述第一张流表指的是所述转发设备中的入口流表。

25 结合第四个方面，在第一种可能的实现方式中，所述元数据中包含端口信息标识区，所述端口信息标识区用于指示所述转发设备的所述至少一个转发端口的信息，所述端口信息标识区中的每个比特位对应指示一个所述转发端口；

30 所述动作域中发送所述数据流的所述至少一个转发端口所述端口信息标识区中指示的所述转发设备的所述至少一个转发端口一一对应。

结合第四个方面的第一种可能的实现方式，在第二种可能的实现方式中，所述处理模块，用于根据所述写元数据信息指示建立所述多路径表的流表项，具体为：

用于查询所述多路径表中是否存在包含所述元数据的流表项；

- 5 若不存在，则用于根据所述写元数据指示，提取所述元数据中所述端口信息标识区指示的所述转发设备的所述至少一个转发端口的信息，对应生成所述多路径表的流表项中的所述动作域。

本发明的第五个方面是提供一种控制器，包括：

- 10 处理器、存储器、总线和通信接口；所述存储器用于存储计算机执行指令，所述处理器与所述存储器通过所述总线连接，当所述控制器运行时，所述处理器执行所述存储器存储的所述计算机执行指令，以使所述控制器执行如第一个方面或第一个方面任意一种可能的实现方式所述的方法。

本发明的第六个方面是提供一种转发设备，包括：

- 15 处理器、存储器、总线和通信接口；所述存储器用于存储计算机执行指令，所述处理器与所述存储器通过所述总线连接，当所述控制器运行时，所述处理器执行所述存储器存储的所述计算机执行指令，以使所述控制器执行如第二个方面或第二个方面任意一种可能的实现方式所述的方法。

- 20 本发明的第七个方面是提供一种多路径转发规则构造系统，包括：第三个方面或第三个方面任意一种可能的实现方式所述的控制器和第四个方面或第四个方面任意一种可能的实现方式所述的转发设备；或者，第五个方面所述的控制器和第六个方面所述的转发设备。

- 25 本实施例提供的多路径转发规则的构造方法、装置及系统，通过控制器调用多路径算法生成数据流的多条路径，之后控制器根据多条路径生成每个转发设备的多端口规则，每个转发设备的多端口规则指示转发设备对应多条路径的至少一个转发端口，控制器根据每个转发设备的多端口规则生成元数据，控制器生成建立流表项的指示建立流表项的指示中包含写元数据指示、跳转到多路径表指示和写元数据指示对应的元数据，控制器将建立流表项的指示发送给每个转发设备，以便每个转发设备依据生成流表项的指示，根据元数据选择转发设备的至少一个转发端口转发所述数据
30 流。由于转发设备根据元数据即可以对端口组合实现复用，因此，控制器

在生成流表项后，不需要查询已经使用过的组，对比历史组的操作桶中的端口组合是否和将建立的组一致，直接将流表项发送给每个转发设备，从而加快了控制器的决策时间。

5 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做一简单地介绍，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性的前提下，还可以根据这些附图获得其他的附图。

图 1 为现有技术多路径转发数据处理流程示意图；

图 2 为本发明实施例提供的一种多路径转发规则构造方法的流程示意图；

图 3 为本发明实施例提供的端口信息标识区示意图；

图 4 为本发明实施例提供的一种多路径转发规则构造方法的流程示意图；

图 5 为本发明实施例提供的另一种多路径转发规则构造方法的流程示意图；

图 6 为本发明实施例提供的另一种多路径转发规则构造方法的流程示意图；

图 7 为本发明实施例提供的另一种多路径转发规则构造方法的流程示意图；

图 8 为本发明实施例多路径示意图；

图 9 为本发明实施例提供的一种转发设备侧多流表示意图；

图 10 为本发明实施例提供的一种控制器的结构示意图；

图 11 为本发明实施例提供的一种转发设备的结构示意图；

图 12 为本发明实施例提供的一种通用网络组件的结构示意图。

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。

- 5 本发明实施例应用场景为典型的SDN网络场景，包含网络控制器和网络转发设备。网络控制器和转发设备间建立数据通道（例如TCP连接），并使用一定的通信标准（如OpenFlow等南北向协议）通信。网络控制器通过对转发设备上的流表（包括普通流表和组表）的操作来实现对数据流转发行为的控制。
- 10 现有技术中，若转发设备采用单张流表来架构数据包处理的管道，对于多个匹配项组合的单流表，该单流表可能有 $M*N$ 条。而如果采用多流表架构数据包处理的管道，则多流表的规模为 $M+N$ 。转发设备会采用多（张）流表的架构来形成数据包处理的管道，可以达到缩小流表空间、增强复杂业务处理、提升处理性能的目标，图1为现有技术多路径转发数据处理流程示意图，
- 15 如图1所示。引入多流表后，数据包在转发设备的处理有可能涉及多个流表的匹配。多个流表的匹配逻辑可能包括：1) 在表1（Flow Table 1）中匹配了数据包头中A、B域的值，在表2（Flow Table 2）中匹配C域的值；2) 在表1（Flow Table 1）中匹配了数据包头中的A、B域，流表项中增加一些新的特征信息存储在元数据（Metadata）中传递给表2（Flow Table 2），在表2（Flow Table 2）
- 20 中匹配元数据（metadata）的信息，并根据匹配的结果执行后续的操作；3) 在表n（Flow Table n）中匹配完成后，跳转到组表（Group Table）的某个表项，进行组表的匹配和对应的操作。需指出，所述表1,表2, 表n以及A、B、C域仅作为标识使用，实际情况可能是不同的表的编号，匹配域的名称和匹配域的数量也可以是各种不同的组合情况。
- 25 控制器需要完成对转发设备的智能管控，其硬件要求是有一定计算能力和网络通信能力的结点，可以部署在台式机、服务器上，亦可部署在虚拟化环境中如虚拟机中。
- 转发设备要求一定的计算能力以处理协议逻辑，同时需要有数据包转发的能力；其处理芯片可能是ASIC，NP或者基于X86的智能网卡芯片。
- 30 本发明实施例在上述场景的基础上，对多流表的组织形式和处理流程进

行了增强和优化，以实现流表项的快速建立。使用多流表中单独的一张表作为多路径表，实现类似组表的功能。该多路径表的匹配域为元数据，需要说明的是，元数据区别于扁平数据，是指数据的若干比特位的组合有特定含义，如每一位对应一个端口号，而扁平的数据必须作为一个整体才有含义。动作

5 域实现组表中的操作桶的功能。元数据中包含转发端口/转发端口组合的信息，元数据中的某一比特位为1代表操作桶中包含该端口，即可能作为多路径转发的端口。下面通过具体的实施例对本发明实施例提供的多路径转发规则构造方法和装置进行说明。

图2为本发明实施例提供的一种多路径转发规则构造方法的流程示意图，本发明实施例的执行主体为控制器。参照图1，该方法包括如下步骤：

10 步骤100、控制器调用多路径算法生成数据流的多条路径。

其中，本实施例对于采用多路径算法不做限定。

步骤101、控制器根据多条路径生成每个转发设备的多端口规则，每个转发设备的多端口规则指示转发设备对应多条路径的至少一个转发端

15 口。

步骤102、控制器根据每个转发设备的多端口规则生成元数据。

具体的，元数据中包含指示转发设备的至少一个转发端口的信息。

步骤103、控制器生成建立流表项的指示。

具体的，建立流表项的指示中包含写元数据指示、跳转到多路径表指

20 示和写元数据指示对应的元数据。

其中，写元数据指示和跳转到多路径表指示跳转到多路径表指示均为操作，写元数据指示用于指示转发设备将写元数据指示对应的元数据写入多路径表中，跳转到多路径表指示用于指示数据包从转发设备的第一张表跳转到多路径表。

步骤104、控制器将建立流表项的指示发送给每个转发设备，以便每个转发设备依据生成流表项的指示，根据元数据选择转发设备的至少一个转发端口转发所述数据流。

25

本实施例提供的多路径转发规则构造方法，通过控制器调用多路径算法生成数据流的多条路径，之后控制器根据多条路径生成每个转发设备的

30 多端口规则，每个转发设备的多端口规则指示转发设备对应多条路径的至

少一个转发端口，控制器根据每个转发设备的多端口规则生成元数据，控制器生成建立流表项的指示建立流表项的指示中包含写元数据指示、跳转到多路径表指示和写元数据指示对应的元数据，控制器将建立流表项的指示发送给每个转发设备，以便每个转发设备依据生成流表项的指示，根据元数据选择转发设备的至少一个转发端口转发所述数据流。由于转发设备根据元数据既可以对端口组合实现复用，因此，控制器在生成流表项后，不需要查询已经使用过的组，对比历史组的操作桶中的端口组合是否和将建立的组一致，直接将流表项发送给每个转发设备，从而加快了控制器的决策时间。

5 优选的，元数据中包含端口信息标识区，端口信息标识区用于指示转发设备的至少一个转发端口的信息，端口信息标识区中的每个比特位对应指示一个转发端口。

具体的，端口信息标识区中的每个比特位对应指示一个多路径转发的端口。图 3 为本发明实施例提供的端口信息标识区示意图，参照图 3，将元数据中的若干比特位如 48 位作为端口信息标识区，当某一比特位置为 1 时，代表该比特位对应的端口将作为多路径转发的端口。如图 3 所示，第 1bit 和第 2bit 位置为 1，代表该表项可以支持第 1 和第二端口多路径转发。当端口较多时，可以利用元数据的部分比特位代表槽位号；亦可以定制化元数据的位数（即扩大比特位数）以适应特殊交换机的需求。

15 在图 2 的基础上，图 4 为本发明实施例提供的一种多路径转发规则构造方法的流程示意图，参照图 4，在步骤 102 之前，还包括：

步骤 105、控制器判断每个转发设备的多端口规则对应的分流比例是否为等比。

具体的，若是等比，则执行图 2 步骤 102。若每个转发设备的多端口规则对应的分流比例不是等比，则控制器使用元数据作为索引，根据元数据定位具有相同分流比例的组表。

需要说明的是，为了保证组的复用，控制器需要为每个转发设备维护一张组表。将元数据作为索引，把所有从相同端口组合的转发的组链接到同一个索引下面。由于每个组可能包含不同的分流比例，例如每个操作桶（action bucket）的权重不一样，不同分流比例的组需要区别对待，若采

用图 2 所示的各个步骤，需要控制器向转发设备发送的流表项中添加分流比例信息，从而增大了元数据实现难度，对于每个组具有的不同分流比例的场景，将元数据作为索引，能够快速查询相同端口组合的组集合，从而快速定位到相同端口组合的组集合，缩小查询的结果空间，从而加速查询时间。

进一步的，在定位具有相同分流比例的组表后，将该组表中的组标识（Group ID）作为流表项发送给转发设备，以使转发设备将该组标识第一张表的动作域。

图5为本发明实施例提供的另一种多路径转发规则构造方法的流程示意图，参照图5，将元数据（metadata）作为索引，例如，图5中的元数据末尾16位为0000000000000011的索引对应的是从1,2端口转发的组的集合，同一个组集合中的组可以通过链表连起来。控制器构造第一张流表规则时，首先根据对应的端口组合构造出元数据（metadata）/索引（index），基于索引（index）查询得到对应的组标识（Group ID），将组标识作为第一张流表中流表项的操作即可。

需要说明的是，上述步骤 105 以及将元数据作为索引的方案、106 与图 2 所示的全部步骤可以分别单独实现，或以组合的形式实现。对于非等比例的分流，要在元数据（metadata）中传递分流的比例或者权重，将加大转发设备识别元数据（metadata）逻辑的复杂度，引入更大的开销。因而，对于既有等比例分流如等价多路径（Equal-Cost Multipath Routing，简称：ECMP），又包含非等比例分流的场景，可以结合采用图 2 所示的方案和和图 5 所示方案。也就是说，对于等比例分流的流表规则，在控制器侧无需存储组表信息，直接下发上述建立流表项的指示表项给转发设备，而对于包含不同分流比例的流表规则，可以采用图 5 所示方案，增加元数据（metadata）索引，以查询到的组标识作为第一张流表规则的动作。此外，不做组的遍历，直接使用新的组标识在这种场景下也是一种可行的方案。

图 6 为本发明实施例提供的另一种多路径转发规则构造方法的流程示意图，本发明实施例的执行主体为转发设备。参照图 6，该方法包括如下步骤：

步骤 200、转发设备接收控制器发送的建立流表项的指示。

具体的，流建立流表项的指示中包含写元数据指示、跳转到多路径表指示和写元数据指示对应的元数据。

步骤 201、转发设备根据建立流表项的指示建立第一张表的流表项，
5 第一张表的流表项中的动作域为写入写元数据指示对应的元数据、以及跳转到多路径表。

其中，第一张流表指的是所述转发设备中的入口流表。写入写元数据指示和跳转到多路径表均为操作，转发设备根据写入写元数据指示将写入写元数据指示对应元数据改写入多路径表中，转发设备根据跳转到多路径
10 表指示将数据包从第一张表跳转到多路径表。

步骤 202、转发设备根据写元数据指示建立多路径表的流表项，多路径表的流表项中，匹配域为元数据，动作域为从元数据对应的至少一个转发端口发送数据流。

本发明实施例提供的多路径转发规则构造方法，通过转发设备接收控
15 制器发送的建立流表项的指示，转发设备根据建立流表项的指示建立第一张表的流表项，第一张表的流表项中的动作域为写入写元数据指示对应的元数据、以及跳转到多路径表，最终转发设备根据写元数据指示建立多路径表的流表项，多路径表的流表项中，匹配域为元数据，动作域为从元数据对应的至少一个转发端口发送数据流。依据该元数据实现了对转发端口
20 的复用，节省了转发设备的存储空间。并且，由于匹配域为元数据，因此，动作域可以包含一系列动作，从而使得各设备商可以针对多路径表进行硬件优化以提高性能，例如优化转发端口的端口信息识别，优化端口选择算法的实现。和传统的 ECMP 不同，转发设备实现组或类似功能的方案，容易扩展新的特性，例如在将数据包转发到某个端口之前增加一系列的动作，
25 使得控制器对转发行为的控制更加灵活。

进一步的，图 6 中所述元数据中包含端口信息标识区，端口信息标识区用于指示转发设备的至少一个转发端口的信息，端口信息标识区中的每个比特位对应指示一个转发端口；

动作域中为根据端口信息标识区中指示的所述转发设备的至少一个
30 转发端口的信息一一对应指示的转发端口作为多路径转发的端口，发送数

据流。

进一步的，在图 6 步骤 202 具体包括；

步骤 202a，转发设备查询多路径表中是否存在包含元数据的流表项。

具体的，若不存在，则转发设备根据写元数据指示，提取元数据中端口信息标识区指示的转发设备的至少一个转发端口的信息，对应生成多路径表的流表项中的动作域。

图 7 为本发明实施例提供的另一种多路径转发规则构造方法的流程示意图，结合控制器和转发设备。参照图 7，对控制器和转发设备之间的交互进行说明，该方法包括如下步骤：

10 步骤 300、控制器调用多路径算法生成数据流的多条路径。

具体的，控制器基于某一种多路径算法生成给定流的多条路径。其中多路径算法可以是任意的一种图算法，例如修改后的 Dijkstra 算法，K-shortest-path 算法等。此外，在获得全网拓扑数据（链路的 UP/DOWN 信息、可用带宽、时延等），以及各种业务流量需求、优先级之后，路由的计算成为了一个规划问题，多路径算法可能使用某种数学规划算法例如 LP（Linear Programming）进行计算。图 8 为本发明实施例多路径示意图，如图 8 所示，假设将流 f 定义为匹配源 IP 地址 192.168.1.0 和目的 IP 地址 192.168.2.0 的所有数据包，多路径算法计算得到的路径为 2 条路径：

20 A->B->E，A->C->E。需指出所得的两条路径只是本例子中多条路径中的一种情况，可能还包含 A->D->E,A->C->B->E 等多种情况，此处仅为描述问题使用，其他情况不一一列举。

步骤 301、控制器根据多条路径生成每个转发设备的多端口规则，每个转发设备的多端口规则指示转发设备对应多条路径的至少一个转发端口。

25 具体的，控制器将步骤1中计算所得结果转化为成每个转发设备的多端口的规则。例如，对于转发设备 A 来说，流 f(SRC IP:192.168.1.0, DST IP:192.168.2.0)对应的转发规则为：

Match field: Source IP address = 192.168.1.0/24，

Destination IP address = 192.168.2.0/24

30 需指出“/24”位的掩码表示一个 IP 地址段的匹配，转发设备即可支持

IP 地址或其他数据包头字段的精确匹配，也可以支持带掩码的通配。由步骤 300 结果可知，控制器对转发设备 A 下发的针对流 f 的流表项应该既可以从端口 1 转发，亦可以从端口 2 转发。

5 步骤 302、控制器根据每个转发设备的多端口规则生成元数据，元数据中包含指示转发设备的至少一个转发端口的信息。

步骤 303、控制器生成建立流表项的指示，建立流表项的指示中包含写元数据指示、跳转到多路径表指示和写元数据指示对应的元数据。控制器将建立流表项的指示发送给每个转发设备

10 具体的，控制器将多端口信息包含在元数据中，构造包含写元数据指示（Write-metadata）和跳转到多路径表指示（Goto multipath table）的建立流表项的指示。具体来说，将元数据中的若干比特位如48位作为端口信息标识区，当某一比特位置为1时，代表该比特位对应的端口将作为多路径转发的端口。如图3所示，第1bit和第2bit位置为1，代表该表项可以支持第一端口和第二端口多路径转发。当端口较多时，可以利用元数据的部分比特位代表槽位
15 号；亦可以定制化元数据的位数（即扩大比特位数）以适应特殊交换机的需求。

具体的，控制器和转发设备间遵循的协议可以包括且不限于 OpenFlow、PCEP、Netconf、IRS、XMPP 等，只要保证控制器和转发设备都能正确解析协议报文，提取出其中的转发规则即可。

20 步骤 304、转发设备接收控制器发送的建立流表项的指示。

转发设备接收到控制器发送的包含转发规则的数据包，进行协议解析，抽取内部的建立流表项的指示，建立流表项的指示中包含写元数据指示、跳转到多路径表指示和写元数据指示对应的所述元数据。

步骤 305、转发设备根据建立流表项的指示建立第一张表的流表项。

25 具体的，第一张表的流表项中的动作域为写入写元数据指示对应的元数据、以及跳转到多路径表，其中，第一张流表指的是转发设备中的入口流表。图 8 为本发明实施例提供的第一张表与多路径表的结构示意图，参照图 8，具体来说，第一张流表指的是包含多流表转发设备中的入口流表，完成第一张流表的匹配后可能伴随有其他流表的匹配。从而实现流水线处
30 理来增强复杂业务的处理能力合并性能力。转发设备建立流表项，对于支

持 OpenFlow 协议的设备，只需将包含两个操作：写元数据指示

(Write-metadata) 和 跳转到多路径表指示 (Goto multipath table) 的流表项存储在转发设备的第一张表 (OpenFlow 流表) 中，该第一张表即可能是硬件流表，也可能是软件流表。对于不支持 OpenFlow 的设备，修改
5 元数据信息指示 (Write-metadata) 和 跳转到多路径表指示 (Goto multipath table) 的逻辑可以内嵌在转发设备的处理逻辑中，只要保证数据包匹配时能按照所述两个逻辑转发即可。

步骤 306、转发设备根据写元数据指示建立多路径表的流表项。

具体的，多路径表的流表项中，匹配域为元数据，动作域为从元数据
10 对应的至少一个转发端口发送数据流。多路径表可以是和组表 (group table) 类似的一个流表，具体来说需要实现组表中的 select 类型 (group type) 的功能，同时至少实现等权重的动作桶 (action buckets) 的功能。本发明的另一种实施例中，多路径表是一种仅做多路径转发的特殊功能表，匹配域为元数据，转发动作完全取决于元数据中包含的端口信息，例
15 如随机从多端口中选择一个端口转发，或者基于包头某些字段的哈希函数决定转发端口。本发明的另一种实施例中，多路径表中的流表规则的匹配域是元数据，对应的动作只要能完成基于某种算法选择元数据中包含端口中特定端口即可，不一定受限于操作桶的实现方式，例如转发设备中的固件中内嵌多路径转发逻辑。

20 步骤 307、转发设备查询多路径表中是否存在包含元数据的流表项。

具体的，当在转发设备在多路径表查询到元数据对应的流表项时，代表可以复用原来的多路径表逻辑，本次流表建立过程结束；如果没有查询到，则执行步骤 308。

步骤 308、转发设备根据写元数据指示，提取元数据中端口信息标识
25 区指示的转发设备的至少一个转发端口的信息，对应生成多路径表的流表项中的动作域。

需要说明的是，动作域中可以包含多个动作桶，每个动作桶对应一个动作，执行多路径转发时，由于需要向多个端口转发数据，因此每一个转发端口对应一个动作桶。

30 具体的，图 9 为本发明实施例提供的一种转发设备侧多流表示意图，如

图 9 所示，其中，左侧为转发设备建立的第一张表，该表的流表项内容为：写元数据指示（Write-metadata）和 跳转到多路径表指示（Goto multipath table），转发设备根据该流表项把相应的数据包转至多路径表（multipath table），即图 8 中右侧的表，该多路径表中，匹配域为元数据，动作域为随机选择元数据包含的端口。例如，图 8 中元数据的第 1bit 和第 16bit 为 1，代表可以选择端口 1 或者端口 16。

图 10 为本发明实施实例提供的一种控制器的结构示意图，该控制器能够执行图 4 所示的各个步骤，参照图 10，该控制器包括：处理模块 10、发送模块 11。

10 处理模块 10，用于调用多路径算法生成数据流的多条路径，根据多条路径生成每个转发设备的多端口规则，每个转发设备的多端口规则指示转发设备对应多条路径的至少一个转发端口，根据每个转发设备的多端口规则生成元数据，元数据中包含指示转发设备的至少一个转发端口的信息，生成建立流表项的指示，建立流表项的指示中包含写元数据指示、跳转到多路径表指示和写元数据指示对应的元数据；

15 发送模块 11，用于将建立流表项的指示发送给每个转发设备，以便每个转发设备依据建立流表项的指示，根据元数据选择转发设备的至少一个转发端口转发数据流。

20 本实施例提供的控制器，通过处理模块调用多路径算法生成数据流的多条路径，根据多条路径生成每个转发设备的多端口规则，每个转发设备的多端口规则指示转发设备对应多条路径的至少一个转发端口，根据每个转发设备的多端口规则生成元数据，元数据中包含指示转发设备的至少一个转发端口的信息，生成建立流表项的指示，建立流表项的指示中包含写元数据指示、跳转到多路径表指示和写元数据指示对应的元数据；发送模块将建立流表项的指示发送给每个转发设备，以便每个转发设备依据建立流表项的指示，根据元数据选择转发设备的至少一个转发端口转发数据流。由于转发设备根据元数据既可以对端口组合实现复用，因此，处理模块在生成流表项后，不需要查询已经使用过的组，对比历史组的操作桶中的端口组合是否和将建立的组一致，直接将流表项发送给每个转发设备，

25 从而加快了控制器的决策时间。

30

优选的，元数据中包含端口信息标识区，端口信息标识区用于指示转发设备的至少一个转发端口的信息，端口信息标识区中的每个比特位对应指示一个转发端口。

5 可选的，所述处理模块 10，还用于在根据每个转发设备的多端口规则生成元数据之前，判断每个转发设备的多端口规则对应的分流比例是否为等比，若是等比，则根据每个转发设备的多端口规则生成所述元数据。

若每个所述转发设备的多端口规则对应的分流比例不是等比，则所述处理模块 10，还用于使用元数据作为索引，根据元数据定位具有相同分流比例的组表。

10 图 11 为本发明实施例提供的一种转发设备的结构示意图，该转发设备能够执行图 6 所示的各个步骤，参照图 11，该转发设备包括：接收模块 20、处理模块 21。

接收模块 20，用于接收控制器发送的建立流表项的指示，建立流表项的指示中包含写元数据指示、跳转到多路径表指示和写元数据指示对应的元数据。
15 元数据。

处理模块 21，用于根据建立流表项的指示建立第一张表的流表项，第一张表的流表项中的动作域为写入写元数据指示对应的元数据、以及跳转到多路径表；还用于根据写元数据指示建立多路径表的流表项，多路径表的流表项中，匹配域为元数据，动作域为从元数据对应的至少一个转发端口发送数据流，第一张流表指的是转发设备中的入口流表。
20 口发送数据流，第一张流表指的是转发设备中的入口流表。

本发明实施例提供的多路径转发规则构造方法，通过接收模块接收控制器发送的建立流表项的指示，建立流表项的指示中包含写元数据指示、跳转到多路径表指示和写元数据指示对应的元数据，处理模块根据建立流表项的指示建立第一张表的流表项，第一张表的流表项中的动作域为写入写元数据指示对应的元数据、以及跳转到多路径表；还用于根据写元数据指示建立多路径表的流表项，多路径表的流表项中，匹配域为元数据，动作域为从元数据对应的至少一个转发端口发送数据流，第一张流表指的是转发设备中的入口流表，依据该元数据实现了对转发端口的复用，节省了转发设备的存储空间。并且，由于匹配域为元数据，因此，动作域可以包含一系列动作，从而使得各设备商可以针对多路径表进行硬件优化以提高
25
30

性能，例如优化转发端口的端口信息识别，优化端口选择算法的实现。和传统的 ECMP 不同，转发设备实现组或类似功能的方案，容易扩展新的特性，例如在将数据包转发到某个端口之前增加一系列的动作，使得控制器对转发行为的控制更加灵活。

- 5 优选的，元数据中包含端口信息标识区，端口信息标识区用于指示转发设备的至少一个转发端口的信息，端口信息标识区中的每个比特位对应指示一个转发端口；

动作域中发送数据流的至少一个转发端口端口信息标识区中指示的转发设备的至少一个转发端口一一对应。

- 10 可选的，所述处理模块 21，处理模块，用于根据所述写元数据信息指示建立所述多路径表的流表项，具体为：

用于查询所述多路径表中是否存在包含所述元数据的流表项；

- 15 若存在，则建立多路径表的流表项完毕；若不存在，则用于根据所述写元数据指示，提取所述元数据中所述端口信息标识区指示的所述转发设备的所述至少一个转发端口的信息，对应生成所述多路径表的流表项中的所述动作域。

- 图 12 为本发明实施实例提供的一种通用网络组件的结构示意图，如图 12 所示为典型的通用网络组件 30，其适用于实施本文本所公开的组件和方法的一项或多项实施例。所述网络组件 30 可以包括处理器 300
20 （可以称为中央处理器单元或 CPU），所述处理器通过总线与包含以下项的存储装置进行通信：辅助存储器 301、只读存储器（ROM）302、随机存取存储器（RAM）303、输入/输出（I/O）装置 304 以及网络连接装置 305。所述处理器 300 可以作为一个或多个 CPU 芯片实施，或者可以为一个或多个专用集成电路（application specific integrated
25 circuit, ASIC）的一部分。

该网络组件 30 可以应用于控制器上，也可以应用在转发设备上。

- 所述辅助存储器 301 通常由一个或多个磁盘驱动器或磁带驱动器组成，且用于对数据进行非易失性存储，而且如果 RAM 303 的容量不足以存储所有工作数据，所述辅助存储器则用作溢流数据存储装置。所
30 述辅助存储器 301 可用于存储程序，所述程序在被选择执行时，被加载

到 RAM 303 中。所述 ROM 302 用于存储在程序执行期间读取的指令以及可能的数据。ROM 302 是非易失性存储装置，其存储容量相对于辅助存储器 301 的较大存储容量而言通常较小。所述 RAM 303 用于存储易失性数据，并且还可能用于存储指令。对 ROM 302 和 RAM 303 两者的存取通常比对辅助存储器 301 的存取快。

当所述装置 30 运行所述存储器中的指令时，所述处理器 300 执行如方法实施例所述的方法步骤，具体的流程请参照方法实施例，这里不再赘述。

本实施例还提供一种多路径转发规则构造系统，该系统包含至少一个图 10 所示的控制器和至少一个图 11 所示转发设备；或者，该系统包含至少一个应用图 12 所示的网络组件的控制器和至少一个应用图 12 所示的网络组件的所示转发设备。相应的，控制器能够执行图 4 所示的各个步骤，并能实现图 4 对应实施例的技术效果；转发设备能够执行图 6 所示的各个步骤，并能实现图 6 对应实施例的技术效果；进一步的，该系统中控制器与转发设备的交互过程如图 7 所示。

本领域普通技术人员可以理解：实现上述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。

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。

权利要求书

1、一种多路径转发规则构造方法，其特征在于，包括：

控制器调用多路径算法生成数据流的多条路径；

5 所述控制器根据所述多条路径生成每个转发设备的多端口规则，每个所述转发设备的多端口规则指示所述转发设备对应所述多条路径的至少一个转发端口；

所述控制器根据每个所述转发设备的多端口规则生成元数据，所述元数据中包含指示所述转发设备的所述至少一个转发端口的信息；

10 所述控制器生成建立流表项的指示，所述建立流表项的指示中包含写元数据指示、跳转到多路径表指示和所述写元数据指示对应的所述元数据；

所述控制器将所述建立流表项的指示发送给每个所述转发设备，以便每个所述转发设备依据所述生成流表项的指示，根据所述元数据选择所述转发设备的所述至少一个转发端口转发所述数据流。

15 2、根据权利要求 1 所述的方法，其特征在于，所述元数据中包含端口信息标识区，所述端口信息标识区用于指示所述转发设备的所述至少一个转发端口的信息，所述端口信息标识区中的每个比特位对应指示一个所述转发端口。

20 3、根据权利要求 1 或 2 所述的方法，其特征在于，所述控制器根据每个所述转发设备的多端口规则生成元数据之前，还包括：

所述控制器判断每个所述转发设备的多端口规则对应的分流比例是否为等比，若是等比，则执行所述控制器根据每个所述转发设备的多端口规则生成所述元数据的步骤。

4、一种多路径转发规则构造方法，其特征在于，包括：

25 转发设备接收控制器发送的建立流表项的指示，所述建立流表项的指示中包含写元数据指示、跳转到多路径表指示和所述写元数据指示对应的所述元数据；

30 所述转发设备根据所述建立流表项的指示建立第一张表的流表项，所述第一张表的流表项中的动作域为写入所述写元数据指示对应的所述元数据、以及跳转到多路径表，其中，所述第一张流表指的是所述转发设备

中的入口流表；

所述转发设备根据所述写元数据指示建立所述多路径表的流表项，所述多路径表的流表项中，匹配域为所述元数据，动作域为从所述元数据对应的至少一个转发端口发送数据流。

- 5 5、根据权利要求4所述的方法，其特征在于，所述元数据中包含端口信息标识区，所述端口信息标识区用于指示所述转发设备的所述至少一个转发端口的信息，所述端口信息标识区中的每个比特位对应指示一个所述转发端口；

10 所述动作域中为根据所述端口信息标识区中指示的所述转发设备的所述至少一个转发端口的信息一一对应指示的所述转发端口作为多路径转发的端口，发送所述数据流。

6、根据权利要求5所述的方法，其特征在于，所述转发设备根据所述写元数据指示建立所述多路径表的流表项，包括；

15 所述转发设备查询所述多路径表中是否存在包含所述元数据的流表项；

若不存在，则所述转发设备根据所述写元数据指示，提取所述元数据中所述端口信息标识区指示的所述转发设备的所述至少一个转发端口的信息，对应生成所述多路径表的流表项中的所述动作域。

7、一种控制器，其特征在于，包括：

20 处理模块，用于调用多路径算法生成数据流的多条路径，根据所述多条路径生成每个转发设备的多端口规则，每个所述转发设备的多端口规则指示所述转发设备对应所述多条路径的至少一个转发端口，根据每个所述转发设备的多端口规则生成元数据，所述元数据中包含指示所述转发设备的所述至少一个转发端口的信息，生成建立流表项的指示，所述建立流表项的指示中包含写元数据指示、跳转到多路径表指示和所述写元数据指示对应的所述元数据；

发送模块，用于将所述建立流表项的指示发送给每个所述转发设备，以便每个所述转发设备依据所述建立流表项的指示，根据所述元数据选择所述转发设备的所述至少一个转发端口转发所述数据流。

30 8、根据权利要求7所述的控制器，其特征在于，所述元数据中包含

端口信息标识区，所述端口信息标识区用于指示所述转发设备的所述至少一个转发端口的信息，所述端口信息标识区中的每个比特位对应指示一个所述转发端口。

9、根据权利要求 7 或 8 所述的控制器，其特征在于，所述处理模块，还用于在根据每个所述转发设备的多端口规则生成元数据之前，判断每个所述转发设备的多端口规则对应的分流比例是否为等比，若是等比，则根据每个所述转发设备的多端口规则生成所述元数据。

10、一种转发设备，其特征在于，包括：

接收模块，用于接收控制器发送的建立流表项的指示，所述建立流表项的指示中包含写元数据指示、跳转到多路径表指示和所述写元数据指示对应的所述元数据；

处理模块，用于根据所述建立流表项的指示建立第一张表的流表项，所述第一张表的流表项中的动作域为写入所述写元数据指示对应的所述元数据、以及跳转到多路径表；还用于根据所述写元数据指示建立所述多路径表的流表项，所述多路径表的流表项中，匹配域为所述元数据，动作域为从所述元数据对应的至少一个转发端口发送数据流，所述第一张流表指的是所述转发设备中的入口流表。

11、根据权利要求 10 所述的转发设备，其特征在于，所述元数据中包含端口信息标识区，所述端口信息标识区用于指示所述转发设备的所述至少一个转发端口的信息，所述端口信息标识区中的每个比特位对应指示一个所述转发端口；

所述动作域中发送所述数据流的所述至少一个转发端口所述端口信息标识区中指示的所述转发设备的所述至少一个转发端口一一对应。

12、根据权利要求 11 所述的转发设备，其特征在于，所述处理模块，用于根据所述写元数据信息指示建立所述多路径表的流表项，具体为：

用于查询所述多路径表中是否存在包含所述元数据的流表项；

若不存在，则用于根据所述写元数据指示，提取所述元数据中所述端口信息标识区指示的所述转发设备的所述至少一个转发端口的信息，对应生成所述多路径表的流表项中的所述动作域。

13、一种控制器，其特征在于，包括：

处理器、存储器、总线和通信接口；所述存储器用于存储计算机执行指令，所述处理器与所述存储器通过所述总线连接，当所述控制器运行时，所述处理器执行所述存储器存储的所述计算机执行指令，以使所述控制器执行如权利要求 1-3 任意一项所述的方法。

5 14、一种转发设备，其特征在于，包括：

处理器、存储器、总线和通信接口；所述存储器用于存储计算机执行指令，所述处理器与所述存储器通过所述总线连接，当所述转发设备运行时，所述处理器执行所述存储器存储的所述计算机执行指令，以使所述转发设备执行如权利要求 4-6 任意一项所述的方法。

10 15、一种多路径转发规则构造系统，其特征在于，包括：权利要求 7-9 任意一项所述的控制器和权利要求 10-12 任意一项所述的转发设备；或者，权利要求 13 所述的控制器和权利要求 14 所述的转发设备。

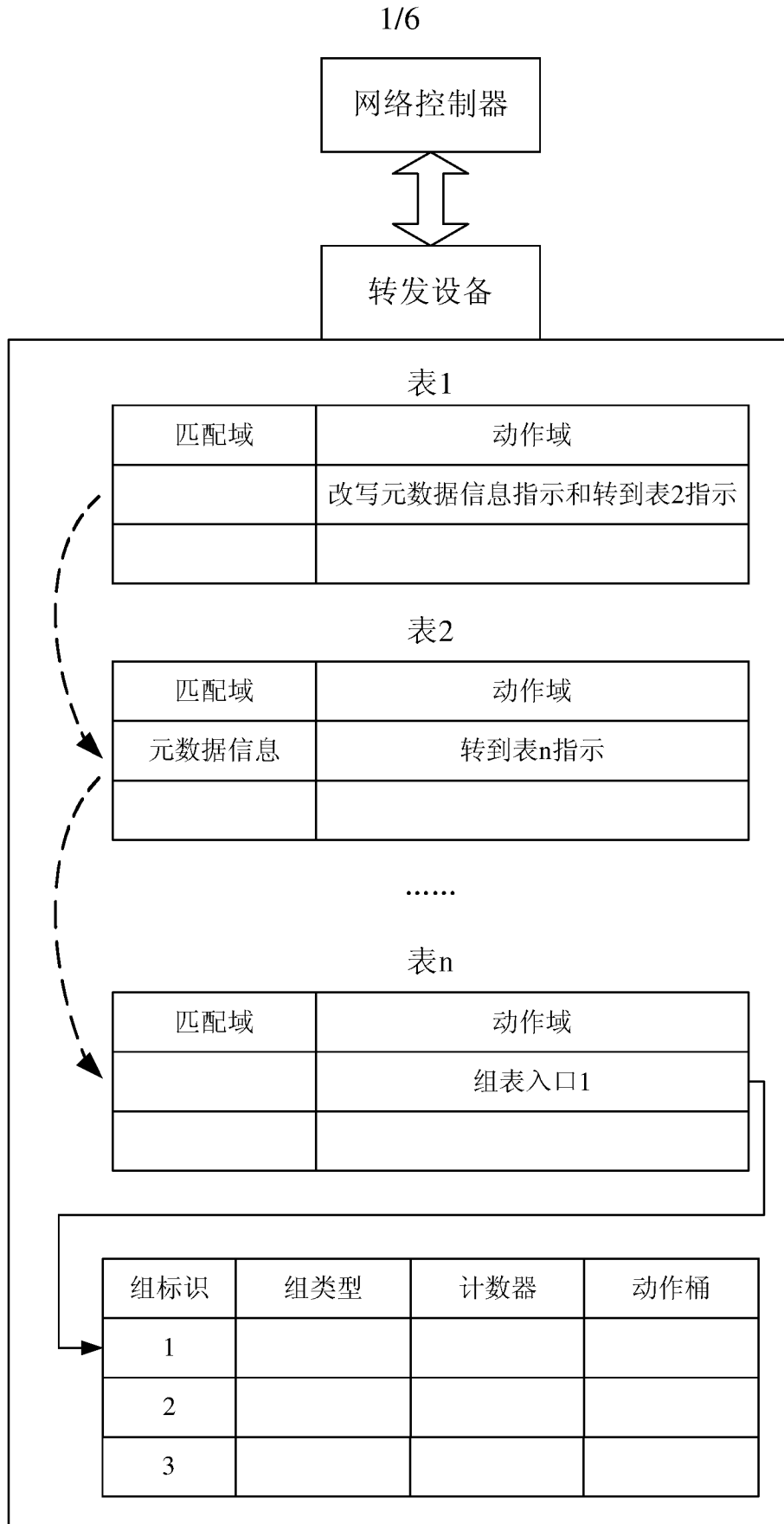


图 1

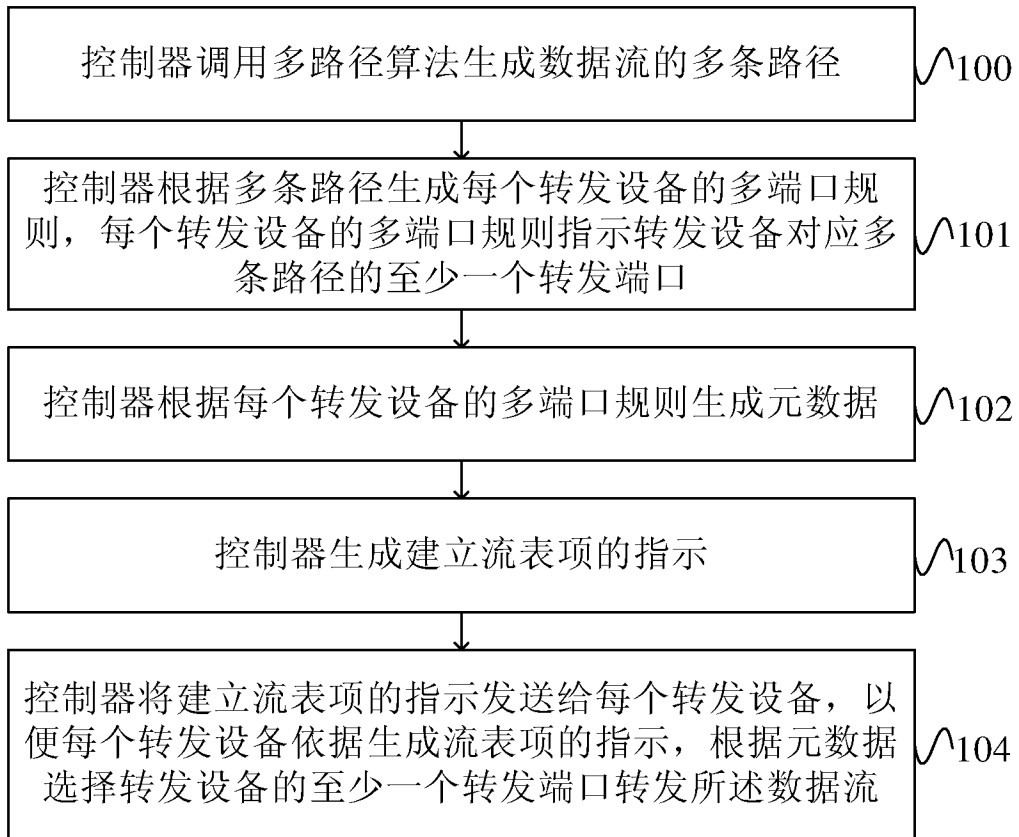


图 2

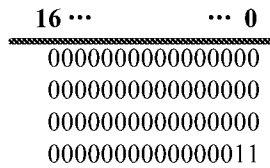


图 3

3/6

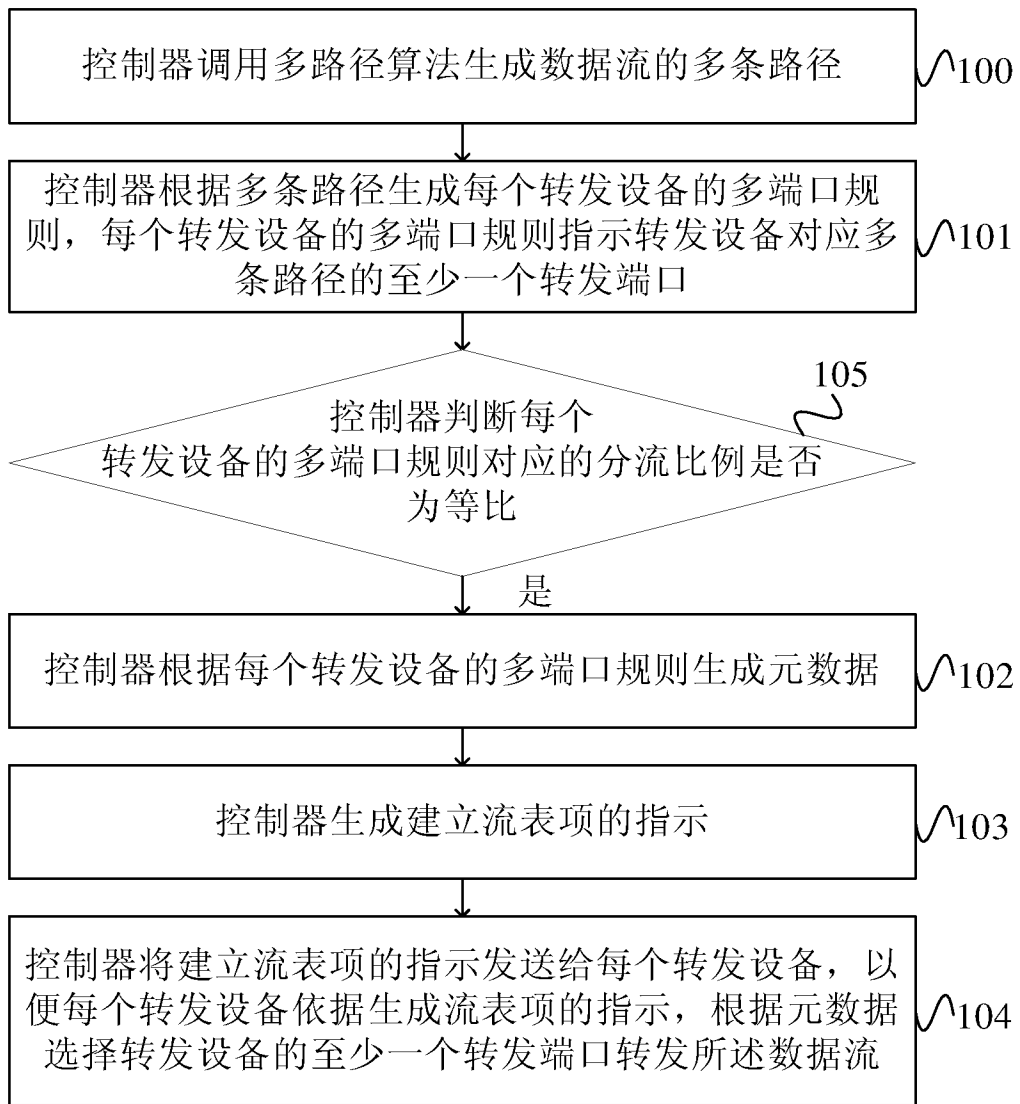


图 4

4/6

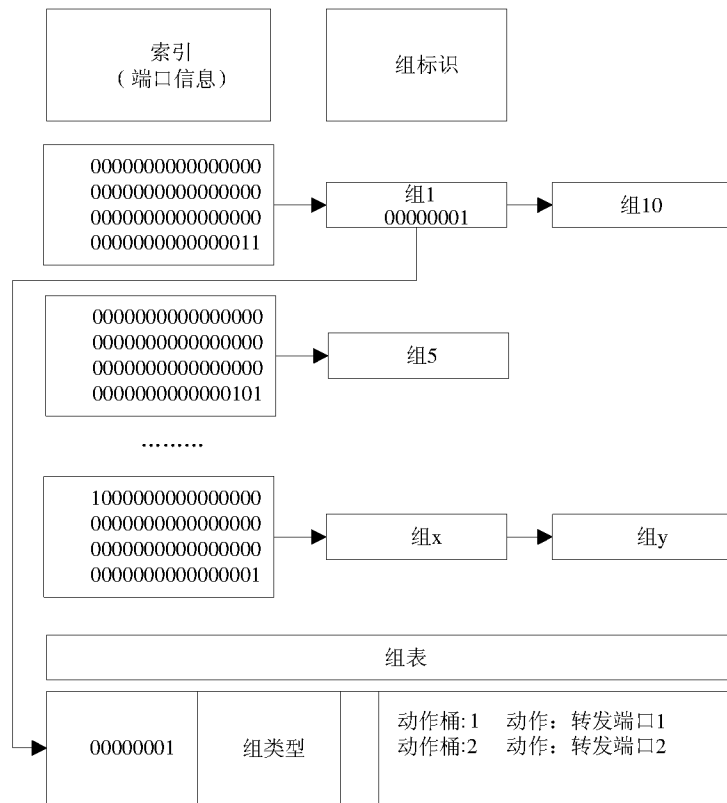


图 5

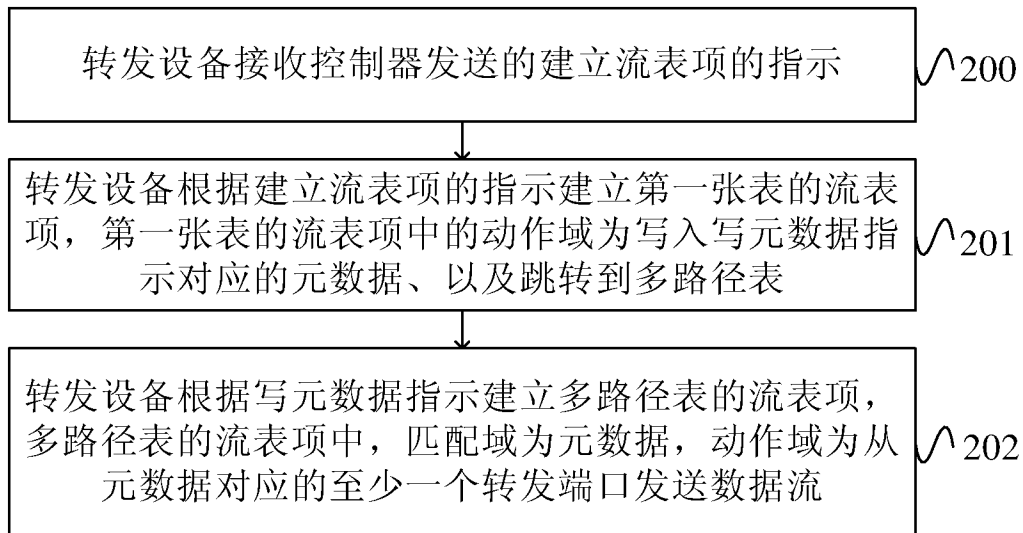


图 6

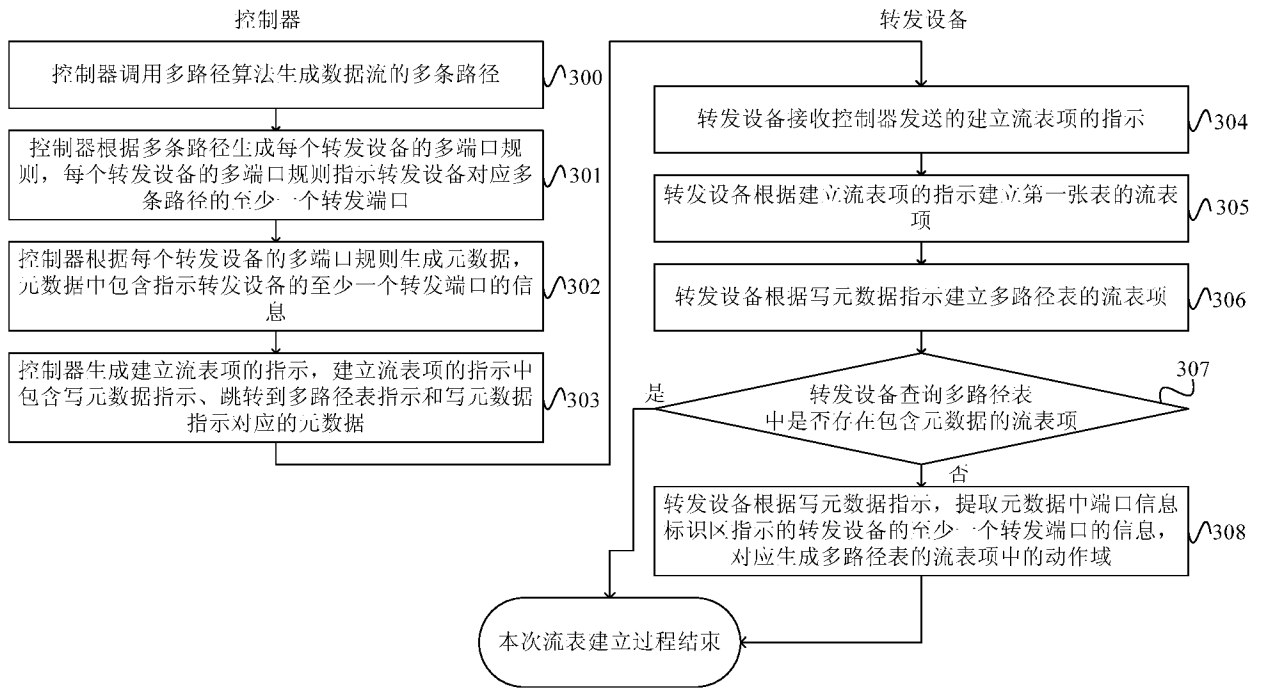


图 7

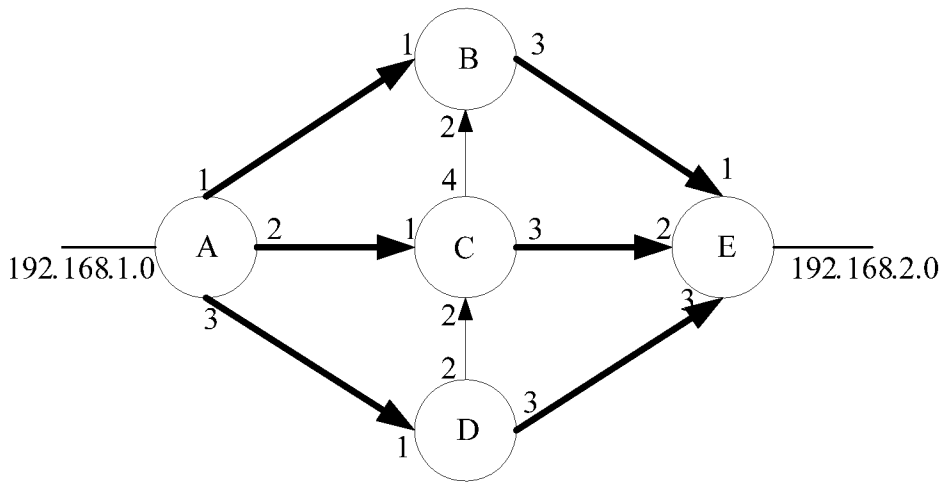


图 8

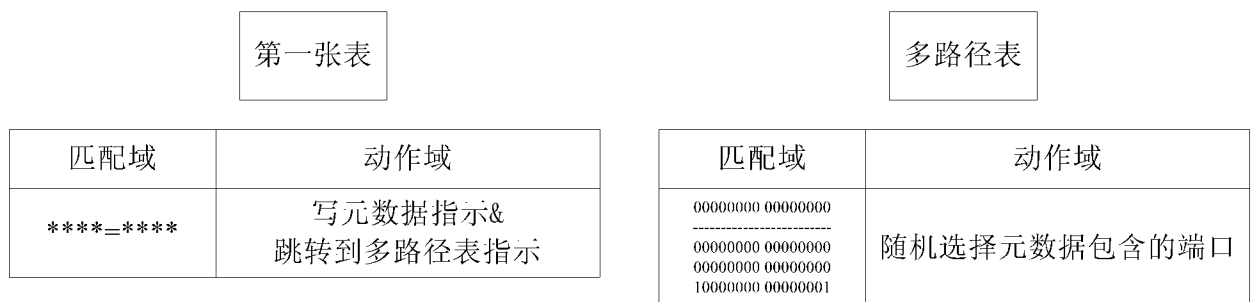


图 9

6/6

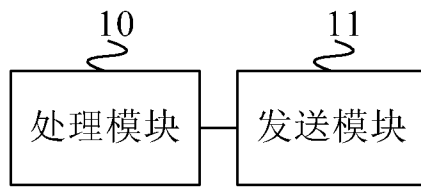


图 10

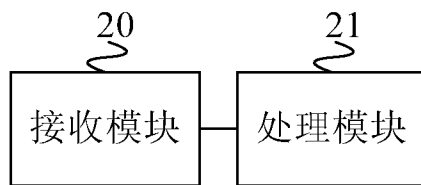


图 11

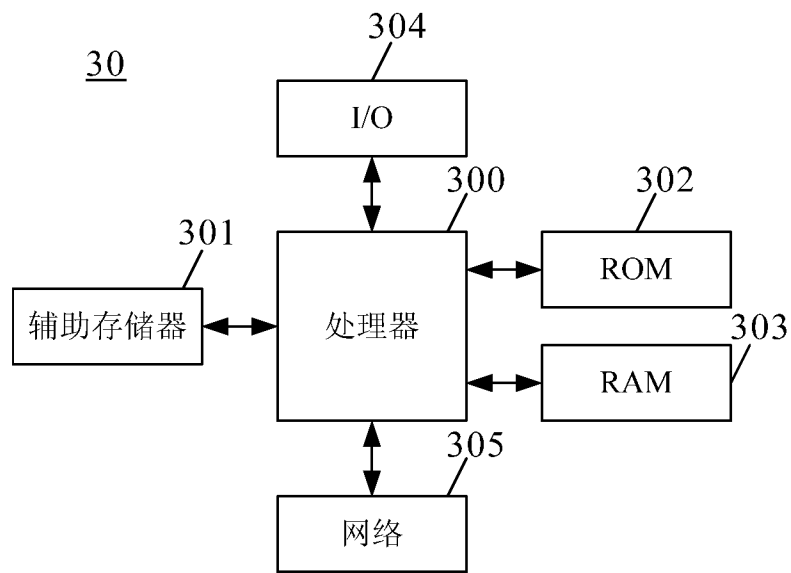


图 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2014/078972

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/74 (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; G06F; H04W; H04Q

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)

CNPAT, WPI, EPODOC, CNKI: key, metadata, bit, port?, multi w rout+, establish+, indicat+, write, controller, SDN, network, Switche?, table, topology, rule, forward, skip, item

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 103329488 A (NIPPON ELECTRIC CO., LTD.) 25 September 2013 (25.09.2013) claims 1 and 7, description, paragraphs [0033]-[0054], and figures 1 and 2	1-15
A	WO 201119792 A2 (MICROSOFT CORPORATION) 29 September 2011 (29.09.2011) the whole document	1-15
A	CN 102957603 A (SHENGKE NETWORK (SOUZHOU) CO., LTD.) 06 March 2013 (06.03.2013) the whole document	1-15
A	CN 103281246 A (HUAWEI TECHNOLOGIES CO., LTD.) 04 September 2013 (04.09.2013) the whole document	1-15
A	CN 103428094 A (HANGZHOU H3C TECHNOLOGIES CO., LTD.) 04 December 2013 (04.12.2013) the whole document	1-15

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

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>
---	---

Date of the actual completion of the international search
07 February 2015

Date of mailing of the international search report
27 February 2015

Name and mailing address of the ISA
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
ZHANG, Liang
Telephone No. (86-10) 62413425

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2014/078972

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103329488 A	25 September 2013	KR 20130121921 A	06 November 2013
		WO 2012096131 A1	19 July 2012
		JP 2014168285 A	11 September 2014
		EP 2665229 A1	20 November 2013
		US 2013279371 A1	24 October 2013
		JP 2014168286 A	11 September 2014
		TW 201246843 A	16 November 2012
		JP 5585667 B	10 September 2014
WO 2011119792 A2	29 September 2011	EP 2550617 B2	30 January 2013
		CN 102804175 A	28 November 2012
		AR 080727 A1	02 May 2012
		US 201123106 A1	
CN 102957603 A	06 March 2013	None	
CN 103281246 A	04 September 2013	None	
CN 103428094 A	04 December 2013	None	

<p>A. 主题的分类</p> <p>H04L 12/741 (2013.01) i</p> <p>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</p>																				
<p>B. 检索领域</p> <p>检索的最低限度文献 (标明分类系统和分类号)</p> <p>H04L; G06F; H04W; H04Q</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNPAT, WPI, EPODOC, CNKI: 元数据, 端口, 比特, 转发, 流表, OpenFlow, 多路径, 表项, 构造, 构建, 规则, 关键, 指示, 写, 跳转, 建立, 控制器, 交换, 软件定义, 组网, key, metadata, bit, port?, multi w rout+, establish+, indicat+, write, controller, SDN, network, Switche?, table, topology, rule</p>																				
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>CN 103329488 A (日本电气株式会社) 2013年 9月 25日 (2013 - 09 - 25) 权利要求1、7, 说明书第[0033]-[0054]段, 图1-2</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>WO 2011119792 A2 (MICROSOFT CORPORATION) 2011年 9月 29日 (2011 - 09 - 29) 全文</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>CN 102957603 A (盛科网络苏州有限公司) 2013年 3月 6日 (2013 - 03 - 06) 全文</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>CN 103281246 A (华为技术有限公司) 2013年 9月 4日 (2013 - 09 - 04) 全文</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>CN 103428094 A (杭州华三通信技术有限公司) 2013年 12月 4日 (2013 - 12 - 04) 全文</td> <td>1-15</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	A	CN 103329488 A (日本电气株式会社) 2013年 9月 25日 (2013 - 09 - 25) 权利要求1、7, 说明书第[0033]-[0054]段, 图1-2	1-15	A	WO 2011119792 A2 (MICROSOFT CORPORATION) 2011年 9月 29日 (2011 - 09 - 29) 全文	1-15	A	CN 102957603 A (盛科网络苏州有限公司) 2013年 3月 6日 (2013 - 03 - 06) 全文	1-15	A	CN 103281246 A (华为技术有限公司) 2013年 9月 4日 (2013 - 09 - 04) 全文	1-15	A	CN 103428094 A (杭州华三通信技术有限公司) 2013年 12月 4日 (2013 - 12 - 04) 全文	1-15
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
A	CN 103329488 A (日本电气株式会社) 2013年 9月 25日 (2013 - 09 - 25) 权利要求1、7, 说明书第[0033]-[0054]段, 图1-2	1-15																		
A	WO 2011119792 A2 (MICROSOFT CORPORATION) 2011年 9月 29日 (2011 - 09 - 29) 全文	1-15																		
A	CN 102957603 A (盛科网络苏州有限公司) 2013年 3月 6日 (2013 - 03 - 06) 全文	1-15																		
A	CN 103281246 A (华为技术有限公司) 2013年 9月 4日 (2013 - 09 - 04) 全文	1-15																		
A	CN 103428094 A (杭州华三通信技术有限公司) 2013年 12月 4日 (2013 - 12 - 04) 全文	1-15																		
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																				
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>																				
<p>国际检索实际完成的日期</p> <p>2015年 2月 7日</p>		<p>国际检索报告邮寄日期</p> <p>2015年 2月 27日</p>																		
<p>ISA/CN的名称和邮寄地址</p> <p>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</p> <p>传真号 (86-10)62019451</p>		<p>受权官员</p> <p>张惊</p> <p>电话号码 (86-10)62413425</p>																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2014/078972

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	103329488	A	2013年 9月 25日	KR	20130121921	A	2013年 11月 6日
				WO	2012096131	A1	2012年 7月 19日
				JP	2014168285	A	2014年 9月 11日
				EP	2665229	A1	2013年 11月 20日
				US	2013279371	A1	2013年 10月 24日
				JP	2014168286	A	2014年 9月 11日
				TW	201246843	A	2012年 11月 16日
				JP	5585667	B	2014年 9月 10日
WO	2011119792	A2	2011年 9月 29日	EP	2550617	A2	2013年 1月 30日
				CN	102804175	A	2012年 11月 28日
				AR	080727	A1	2012年 5月 2日
				US	2011239106	A1	2011年 9月 29日
CN	102957603	A	2013年 3月 6日	无			
CN	103281246	A	2013年 9月 4日	无			
CN	103428094	A	2013年 12月 4日	无			