(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 104580026 B
(45)授权公告日 2019.02.15

(21)申请号 201510082324.9

(22)申请日 2011.07.21

(65)同一申请的已公布的文献号
申请公布号 CN 104580026 A

(43)申请公布日 2015.04.29

(30)优先权数据
2010-200690 2010.09.08 JP

(62)分案原申请数据
201180043295.5 2011.07.21

(73)专利权人 日本电气株式会社
地址 日本东京都

(72)发明人 铃木洋司 高岛正德 久保田一志 伊泽彻 林将志

(74)专利代理机构 中科专利商标代理有限责任公司 11021
代理人 闫晔

(51)Int.Cl.
H04L 12/937(2013.01)

(56)对比文件
US 2004202158 A1,2004.10.14,
戴锦友等.电信级以太网虚拟硬件方法的研究.《算机科学》.2010,第37卷(第2期),

审查员 刘莉

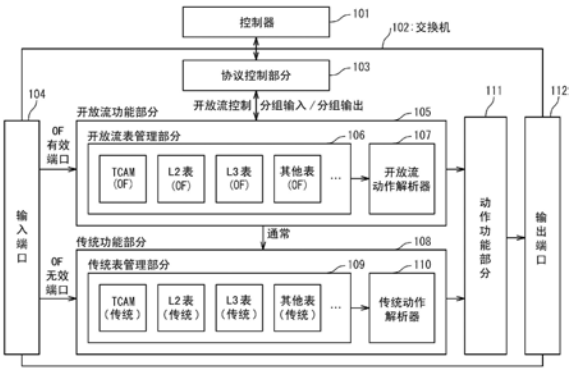
权利要求书1页 说明书16页 附图15页

(54)发明名称

交换系统、交换控制系统和存储介质

(57)摘要

一种交换机系统,通过使用交换机中的表作为现有资源,实现开放流表的条目数量的扩展。具体地,交换机基于在每个所述表中定义的条件和处理内容通过逻辑上合并多个表来配置开放流表,所述多个表中的每一个表定义了给定分组的处理。交换机查阅开放流表,以确定对接收分组的处理内容。交换机基于所确定的处理内容执行接收分组的处理。



CN 104580026 B

1.一种用于转发分组的通信装置,所述通信装置包括:

第一单元,用于基于预定控制协议与控制器进行通信,所述控制器能够向多个通信装置发送第一分组转发规则;以及

第二单元,用于基于所述分组的输入端口,确定根据用于存储所述第一分组转发规则的第一区域还是用于存储由所述通信装置设置的第二分组转发规则的第二区域来处理分组。

2.根据权利要求1所述的通信装置,

其中,所述第一单元从所述控制器接收所述第一分组转发规则,所述第一分组转发规则包括:用于基于所述分组中包括的多个信息来识别所述分组的识别规则和与所述识别规则相对应的分组处理规则。

3.根据权利要求1所述的通信装置,

其中,所述第二单元基于所述输入端口是否与所述控制协议相关联,确定根据所述第一区域还是所述第二区域来处理所述分组。

4.一种通信方法,包括:

用于转发分组的通信装置基于预定控制协议与控制器进行通信,所述控制器能够向多个通信装置发送第一分组转发规则;以及

基于所述分组的输入端口,确定根据用于存储所述第一分组转发规则的第一区域还是用于存储由所述通信装置设置的第二分组转发规则的第二区域来处理所述分组。

交换系统、交换控制系统和存储介质

[0001] 本申请是2013年3月8日提交的、申请号为201180043295.5、发明名称为“交换系统、交换控制系统和存储介质”的专利申请的分案申请。

技术领域

[0002] 本发明涉及交换系统，并具体地涉及其中每个交换机具有多个表的交换系统。

背景技术

[0003] 为了控制网络系统中的通信路径，最近已经开发出来采用开放流(OpenFlow)技术作为针对通信设备的控制协议的路径控制方法。根据开放流技术控制路径的网络称为开放流网络。

[0004] 在开放流网络中，控制器(例如OFC(开放流控制器))操作交换机(例如OFS(开放流交换机))的每个开放流表，以控制交换机的行为。控制器通过安全信道(Secure Channel)与交换机相连，以通过使用符合开放流协议的控制消息来控制交换机。

[0005] 开放流网络中的交换机是构成开放流网络的边缘交换机和核心交换机，并且交换机是在控制器的控制之下。从在开放流网络中的输入侧边缘交换机处接收分组到在开放流网络中的输出侧边缘交换机处发送分组的一些列类型的分组处理称为流。

[0006] 开放流表是其中注册有流条目的表，其定义了对符合预定匹配条件(规则)的分组(通信数据)要执行的预定处理内容(动作)。

[0007] 根据包括在分组中协议层级的每层的首部字段中的目的地址、源地址、目的端口和源端口中的一些或所有的各种组合，定义流条目的规则，并且其可以被区别。上述地址包括MAC地址(媒体访问控制地址)和IP地址(因特网协议地址)。此外，与进入端口有关的信息也可用作流条目的规则。

[0008] 流条目的动作指示操作，例如“向特定的端口输出”、“丢弃”和“重写首部”。例如，当输出端口的标识信息(例如输出端口号)出现在流条目的动作中时，交换机向该端口输出分组，并且当输出端口的标识信息没有出现时，交换机丢弃该分组。备选地，当首部信息出现在流条目的动作中时，交换机基于首部信息重写分组的首部。

[0009] 开放流网络中的交换机对符合流条目的规则的分组组(分组序列)执行流条目的动作。

[0010] 开放流交换机的细节在非专利文献1和2中描述。

[0011] 需要大容量开放流表来控制网络上大量的流。在当前情况下，用于开放流表的TCAM(三重内容可寻址存储器)不具有大容量，并且因此不能保证必须和充足的容量。此外，难以增加在开放流表中使用的交换机的每个表(主要是TCAM)自身的容量。

[0012] 作为解决上述问题的方法之一，可以使用外部TCAM，但这花费开销。此外，在用于高速传输(例如10G多端口)(具有多端口的网络设备，其可以与10G比特/秒的数据传递速率相对应)的设备中，不能使用外部TCAM。现在，不存在能够至少在10G交换机中操作的外部TCAM。

[0013] 引用列表

[0014] [非专利文献1]“The OpenFlow Switch Consortium”<http://www.openflowswitch.org/>

[0015] [非专利文献2]“OpenFlow Switch Specification Version 0.9.0(Wire Protocol 0x98)July 20,2009

[0016] 当前维护者:Brandon Heller(brandonh@stanford.edu)”<http://www.openflowswitch.org/documents/openflow-spec-v0.9.0.pdf>

发明内容

[0017] 本发明的目的是通过使用作为现有资源的交换机中的表,实现开放流表的条目数量的扩展。

[0018] 本发明的交换系统包括:开放流功能部分,通过基于在每个表中定义的条件和处理内容而在逻辑上合并多个表来配置开放流表,多个表中的每个表定义对预定接收分组的处理,并且开放流功能部分查阅开放流表以确定对接收分组的处理内容;以及动作功能部分,基于预定处理内容对接收分组执行处理。

[0019] 在开放流交换机中执行的本发明的交换控制方法,该交换控制方法包括:通过基于在每个表中定义的条件和处理内容在逻辑上合并多个表来配置开放流表,多个表中的每个表定义对预定接收分组的处理;查阅开放流表以确定对接收分组的处理内容;以及基于预定处理内容执行接收分组的处理。

[0020] 根据本发明的程序是使用作交换机的计算机执行上述交换控制方法中的处理的程序。应当注意,根据本发明的程序可以存储在存储单元和存储介质中。

[0021] 本发明允许控制器能够使用交换机中的多个表作为一个大容量的开放流表。

附图说明

[0022] 图1是示出了根据本发明的交换系统的配置示例的概念性示意图;

[0023] 图2是示出了根据本发明的第一示例性实施例的开放流功能部分的细节的概念性示意图;

[0024] 图3是示出了根据本发明的第一示例性实施例的开放流动作解析器的细节的概念性示意图;

[0025] 图4是示出了开放流表控制的概述的示意图;

[0026] 图5是示出了第一类型的开放流表控制的示例的细节的示意图;

[0027] 图6是示出了第二类型的开放流表控制的示例的细节的示意图;

[0028] 图7A是示出了分组流入情况下,交换系统的第一操作示例的示意图;

[0029] 图7B是示出了分组流入情况下,交换系统的第一操作示例的示意图;

[0030] 图8A是示出了分组流入情况下,交换系统的第二操作示例的示意图;

[0031] 图8B是示出了分组流入情况下,交换系统的第二操作示例的示意图;

[0032] 图9是示出了第一类型的开放流表控制的特定示例的示意图;

[0033] 图10是示出了开放流表控制的第二方法的特定示例的示意图;

[0034] 图11是示出了根据本发明的第二示例性实施例的交换系统的示例1的细节的概念

性示意图；

[0035] 图12是示出了根据本发明的第二示例性实施例的交换系统的示例2的细节的概念性示意图；以及

[0036] 图13是示出了根据本发明的第二示例性实施例的交换系统的示例3的细节的概念性示意图。

具体实施方式

[0037] [第一示例性实施例]

[0038] 在下文中将参考附图描述本发明的第一示例性实施例。

[0039] (系统结构)

[0040] 如图1所示，根据本发明的交换系统包括控制器101和交换机102。

[0041] 控制器101在符合开放流协议的处理中控制交换机102。

[0042] (交换机结构)

[0043] 交换机102包括协议控制部分103、输入端口104、开放流功能部分105、传统功能部分108、动作功能部分111和输出端口112。

[0044] 当控制器101执行用于在符合开放流协议的处理中控制交换机102的通信时，协议控制部分103执行控制器101和交换机102之间的协议控制。协议控制部分103不需要提供在交换机102中，而可以提供在交换机102之前的级中。

[0045] 输入端口104是分组输入接口。输入端口104具有开放流有效端口和开放流无效端口。开放流有效端口是符合开放流协议的输入端口，而开放流无效端口是不符合开放流协议的输入端口。

[0046] 开放流功能部分105对从开放流有效端口输入的分组执行处理。

[0047] 开放流功能部分105包括开放流表管理部分106和开放流动作解析器107。

[0048] 开放流表管理部分106保留交换机102使用的开放流表。在开放流表中定义针对符合开放流协议的分组的动作(开放流处理的动作)。

[0049] 开放流动作解析器107基于开放流表管理部分106的查询结果，确定开放流处理的动作。

[0050] 传统功能部分108对从开放流无效端口输入的分组执行处理。

[0051] 传统功能部分108包括传统表管理部分109和传统动作解析器110。

[0052] 传统表管理部分109是交换机102使用的传统表。传统表管理部分109定义了针对不符合开放流协议的分组(通常分组等)的动作(传统处理的动作)。

[0053] 传统动作解析器110基于传统表管理部分109的查询结果，确定传统处理的动作。在传统处理中，使用通常的交换功能。

[0054] 动作功能部分111执行在开放流功能部分105或传统功能部分108中确定的动作。

[0055] 输出端口112是分组输出接口。

[0056] (开放流处理和传统处理之间的区别)

[0057] 在开放流处理中，通过外部控制器控制分组路径。控制器选择整个网络中的最优路径。相反，在传统处理中，如在通常的交换机和路由器中一样，通过自主分配来控制路径。通常的交换机和路由器从与它们的环境有关的信息确定网络状态，以选择最优路径。

[0058] 在开放流处理中，可以基于多达12种类型的信息的组合来标识分组。另一方面，在传统处理中，用于标识分组的信息的类型个数少，例如在L2网络情况下的目的MAC地址和在L3网络情况下的目的IP地址。针对此原因，难以执行良好的流控制。例如，在传统处理中，具有相同目的IP地址但不同源TCP端口号的流被确定为不同的流，从而选择不同的路径。

[0059] (交换系统的完整操作)

[0060] 在下文中将描述图1中的交换系统的完整操作。

[0061] (分组输入)

[0062] 当新分组流入交换机102中时，交换机102在输入端口104处接收该分组。

[0063] 交换机102检查接收该分组的输入端口104是否是开放流有效端口。例如，交换机102通过查阅交换机102自身或输入端口104的设置信息(config)，检查接收该分组的输入端口104是否是开放流有效端口。

[0064] (从分组输入到开放流处理的转换)

[0065] 当输入端口是开放流有效端口时，交换机102将分组从输入端口104向开放流功能部分105传递。

[0066] (开放流处理)

[0067] 开放流功能部分105在保留交换机102的多个表的开放流表管理部分106中对传递的分组执行查询处理。

[0068] 然后，开放流功能部分105基于查询结果和开放流动作解析器107中每个表的优先级，确定该分组的动作。优先级可以称为优先等级。

[0069] (从开放流处理到动作的执行的转换)

[0070] 当所确定的动作是“分组输入”(向控制器询问分组的动作)时(例如当没有流条目存在时，不能确定动作)，开放流功能部分105通过协议控制部分103向控制器101发出询问(例如分组的传送)。初始地，可以将除了目标是传统处理的分组的所有分组的动作设置为“分组输入”，而不加任何条件。开放流功能部分105接收作为对询问的响应的“分组输出”(来自控制器的对行为询问的结果)，确定作为分组动作的内容，并将其注册到开放流表管理部分106保留的表中。在下文中，开放流功能部分105确定遵循与在开放流动作解析器107中的上述分组相同规则的分组的动作。

[0071] 开放流功能部分105基于所确定的动作向动作功能部分111传送该分组。即，处理该分组的主单元从开放流功能部分105转换到动作功能部分111。

[0072] (从开放流处理到传统处理的转换)

[0073] 当所确定的动作是“通常”(使用传统功能部分108处理分组)时，开放流功能部分105向传统功能部分108传递该分组。即，处理该分组的主单元从开放流功能部分105转换到传统功能部分108。

[0074] (从分组输入到传统处理的转换)

[0075] 当输入端口104是开放流无效端口或先前(之前)在开放流功能部分105中确定的分组的动作是“通常”时，交换机102将该分组从输入端口104向传统功能部分108传递。即，处理该分组的主单元从输入端口104转换到传统功能部分108。

[0076] (传统处理)

[0077] 传统功能部分108在被配置有交换机102的多个表的传统表管理部分109中对接收

的分组执行查询处理。

[0078] (从传统处理到动作的执行的转换)

[0079] 然后,传统功能部分108基于查询结果和传统动作解析器110中每个表的优先级,确定该分组的动作。即,处理该分组的主单元从传统功能部分108转换到动作功能部分111。

[0080] 传统功能部分108到传统动作解析器110的传统处理使用通常的交换功能,因此省略其详细描述。

[0081] (动作的执行)

[0082] 动作功能部分111执行在开放流功能部分105或传统功能部分108中确定的对该分组的动作。

[0083] 作为开放流功能部分105中确定的动作的示例,列举了重写首部信息、从指定输出端口输出分组和丢弃分组的示例。作为传统功能部分108中确定的动作的示例,通过路由来传送分组等。然而,本发明不限于这些示例。

[0084] 最后,当要执行的动作包括“分组输出”时,动作功能部分111根据该动作的内容,从合适的输出端口112输出分组。

[0085] (通过控制器控制开放流表)

[0086] 控制器101可以通过协议控制部分103控制交换机102的开放流表管理部分106。这里,“控制开放流表管理部分106”的含义是开放流表中流条目的注册/改变/删除/批量删除等。

[0087] 保留在开放流表管理部分106中用于开放流表的交换机102的每个表不一定符合开放流规范中定义的所有操作。

[0088] 针对此原因,考虑到开放流表管理部分106针对开放流表保留的各个表可以实现的功能(可以设置的动作),控制器101需要控制开放流表管理部分106。

[0089] (开放流功能部分的细节)

[0090] 图2是示出了本发明的开放流功能部分105的细节的示意图。

[0091] 开放流功能部分105、开放流表管理部分106、开放流动作解析器107和动作功能部分111与其在图1中具有相同的机制和功能。

[0092] 开放流功能部分105包括开放流表管理部分106和开放流动作解析器107。

[0093] 开放流表管理部分106包括表组113和查询功能部分114。

[0094] 表组113是构成开放流表的表组。

[0095] 查询功能部分114基于表组113查询输入分组的数据。

[0096] 查询功能部分114包括L2/L3/其他表(OF)查询功能部分115和TCAM(OF)查询功能部分116。

[0097] “OF”是“开放流”的缩写。

[0098] L2/L3/其他表(OF)查询功能部分115针对输入分组查阅L2表(OF)、L3表(OF)和其他表(OF),以查询条目。多播路由表作为其他表(OF)的示例。即,L2/L3/其他表(OF)查询功能部分115以协议的单元针对输入分组而查询表。

[0099] TCAM(OF)查询功能部分116针对输入分组查阅TCMA(OF),以查询条目。即,TCAM(OF)查询功能部分116针对输入分组查询TCAM。

[0100] (开放流功能部分的操作)

[0101] 在下文中将描述图2中示出的开放流功能部分105的操作。

[0102] 输入端口104将从开放流有效端口输入的分组向开放流功能部分105传递。

[0103] 开放流功能部分105在被配置有交换机102的多个表的开放流表管理部分106中对传递的分组执行查询处理。

[0104] 此时，开放流表管理部分106的查询功能部分114基于构成开放流表的表组113中注册的条目信息，执行查询处理。

[0105] 具体地，在查询功能部分114中，L2/L3/其他表(OF)查询功能部分115首先执行查询处理，然后TCAM(OF)查询功能部分116执行查询处理。

[0106] 查询功能部分114向开放流动作解析器107传递查询结果。

[0107] 开放流动作解析器107基于查询结果和每个表的优先级，确定分组的动作。

[0108] (硬件的示例)

[0109] 在下文中将描述用于实现根据本发明的交换系统的硬件的特定示例。

[0110] 作为控制器101的示例，列举了计算机的示例，例如PC(个人计算机)、工作站、主机和超级计算机。控制器101可以是安装在计算机上的扩展板或构建在实体机器上的虚拟机(虚拟机(VM))。

[0111] 作为交换机102的示例，列举了以下示例：L3交换机(层3交换机)、L4交换机(层4交换机)、L7交换机/应用交换机(层7交换机)或网络交换机(网络交换机)(例如多层交换机)。此外，作为交换机102的示例，列举了以下示例：路由器(路由器)、代理(代理)、网关、防火墙、负荷均衡器(负荷分配设备)、频带控制器/安全监控器控制器(看门人)、基站、接入点(AP)、通信卫星(CS)和具有多个通信端口的计算机。

[0112] 可以通过基于程序运行并执行预定处理的处理器以及存储该程序和各种类型数据的存储器，实现协议控制部分103、开放流功能部分105、开放流表管理部分106、开放流动作解析器107、传统功能部分108、传统表管理部分109、传统动作解析器110和动作功能部分111。

[0113] 作为上述处理器的示例，列举了如下示例：CPU(中央处理单元)、网络处理器(NP)、微处理器、微控制器或具有专用功能的半导体集成电路(集成电路(IC))。

[0114] 作为上述存储器的示例，列举了如下示例：RAM(随机存取存储器)、半导体存储器件(例如ROM(只读存储器))、EEPROM(电可擦写可编程只读存储器)和闪存、辅助存储器件(例如HDD(硬盘驱动器)和SSD(固态驱动器))、可移除盘(例如DVD(数字通用盘))和存储介质(介质)(例如SD存储卡(安全数字存储卡))。可以采用缓冲器和寄存器。备选地，可以采用使用DAS(直接附接存储)、FC-SAN(光纤信道-存储区域网络)、NAS(网络附接存储)、IP-SAN(IP-存储区域网络)等。

[0115] 可以集成处理器和存储器。例如，近年来，微计算机等已经集成到一个芯片中。因此，安装在电子设备中的单芯片微计算机可以具有处理器和存储器。

[0116] 备选地，协议控制部分103、开放流功能部分105、开放流表管理部分106、开放流动作解析器107、传统功能部分108、传统表管理部分109、传统动作解析器110和动作功能部分111中的每个可以是安装在计算机上的扩展板或构建在实体机器上的虚拟机(VM)。

[0117] 作为输入端口104和输出端口112的示例，列举了以下示例：半导体集成电路(例如符合网络通信的板(主板或I/O板))、网络适配器(例如NIC(网络接口卡)或类似扩展板)、通

信设备(例如天线)以及通信端口(例如连接端口(连接器))。

[0118] 此外,作为输入端口104和输出端口112使用的网络的示例,列举了以下示例:因特网、LAN(局域网)、无线LAN(无线LAN)、WAN(广域网)、骨干网(骨干网)、有线电视(CATV)线、固定电话网、移动电话网、WiMAX(IEEE 802.16a)、3G(第三代移动通信)、租赁线(租赁线)、IrDA(红外数据协会)、蓝牙(注册商标)、串行通信线以及数据总线。

[0119] 应当注意,协议控制部分103、开放流功能部分105、开放流表管理部分106、开放流动作解析器107、传统功能部分108、传统表管理部分109、传统动作解析器110和动作功能部分111中的每个可以是模块(模块)、组件(组件)或专用设备,或它们的起始(调用)程序。

[0120] 然而,本发明不限于这些示例。

[0121] (开放流动作解析器的细节)

[0122] 图3是示出了本发明的开放流动作解析器107的细节的示意图。

[0123] 开放流动作解析器107和TCAM(OF)查询功能部分116与其在图2中具有相同的机制和功能。

[0124] 通过调整TCAM(OF)的条目映射,开放流动作解析器107可以实现为TCAM(OF)查询功能部分116的一部分。

[0125] 为此原因,开放流动作解析器107和TCAM(OF)查询功能部分116实质上构成一个功能块(TCAM(OF)查询&开放流动作解析器)。该功能块具有表间优先级117和TCAM(OF)中的条目118。

[0126] 表间优先级117指示期望动作优先级。TCAM(OF)中的条目118指示与优先级对应的TCAM(OF)中的条目映射。

[0127] TCAM(OF)中的条目118包括TCAM(OF)查询条目组119、L2表(OF)查询结果查阅条目120、L3表(OF)查询结果查阅条目121、其他表(OF)查询结果查阅条目122以及未命中(Miss-hit)条目123。

[0128] TCAM(OF)查询条目组119是用于实现TCAM(OF)查询功能部分116中的TCAM(OF)查询的条目集合。L2表(OF)查询结果查阅条目120是用于查阅L2表(OF)的查询结果的条目。L3表(OF)查询结果查阅条目121是用于查阅L3表(OF)的查询结果的条目。其他表(OF)查询结果查阅条目122是用于查阅其他表(OF)的查询结果的条目。未命中条目123是用于处理没有命中任何条目的作为“未命中”的条目。即,这些是定义上述相应动作的条目。

[0129] (开放流动作解析器的操作)

[0130] 在下文中将描述图3中的开放流动作解析器107的操作。

[0131] 作为开放流表中的查询顺序,最后,TCAM(OF)查询功能部分116执行查询。应当注意的是,在TCAM(OF)查询功能部分116的查询中,开放流动作解析器107可以通过调整可以查阅L2/L3/其他表的查询结果的交换机中的TCAM(OF)的条目映射来实现。

[0132] 例如,如在表间优先级117中的针对每个表的期望动作优先级可以通过如TCAM(OF)中的条目118中的映射来处理。

[0133] 在图3所示的示例的情况中,表间优先级117变成从最高优先级“全兼容(与TCAM(OF)等同)”、“L2兼容(与L2表(OF)等同)”、“L3兼容(与L3表(OF)等同)”、“其他兼容(与其他表等同)”。

[0134] 用作TCAM(OF)中的条目映射,开放流动作解析器107可以按照从TCAM(OF)的最高

查询优先级的如下顺序来布置“TCAM(OF)查询功能实现条目组119”、“L2表(OF)查询结果查阅条目120”、“L3表(OF)查询结果查阅条目121”、“其他表(OF)查询结果查阅条目122”和“未命中条目123”。

[0135] (TCAM(OF)查询条目组)

[0136] TCAM(OF)查询功能实现条目组119是用于实现针对查询功能的TCAM(OF)查询功能部分116的条目组。当输入分组命中TCAM(OF)查询功能实现条目组119的任意条目时，开放流动作解析器107选择该条目的动作，作为针对该分组的动作。

[0137] (L2表(OF)查询结果查阅条目)

[0138] 当与输入分组对应的条目在L2表(OF)上存在时，基于L2表(OF)之前的查询结果确定命中了L2表(OF)查询结果查阅条目120。

[0139] 例如，当与输入分组对应的条目在L2表(OF)的查询中存在时，设置标记“X＝1”，并且当已经设置标记“X＝1”时，确定在TCAM(OF)的L2表(OF)查询结果查阅条目120中命中了该条目。

[0140] 当确定L2表(OF)查询结果查阅条目120被命中时，开放流动作解析器107选择L2表(OF)的条目的动作作为针对该分组的动作。

[0141] (L3表(OF)查询结果查阅条目)

[0142] 当与输入分组对应的条目在L3表(OF)上存在时，基于L3表(OF)之前的查询结果，确定命中了查询结果查阅条目121。

[0143] 例如，当与输入分组对应的条目在L3表(OF)的查询中存在时，设置标记“Y＝1”，并且当已经设置了标记“Y＝1”时，确定在TCAM(OF)上的L3表(OF)查询结果查阅条目121中命中了条目。

[0144] 当确定L3表(OF)查询结果查阅条目121被命中时，开放流动作解析器107选择L3表(OF)的条目的动作作为针对该分组的动作。

[0145] (其他表(OF)查询结果查阅条目)

[0146] 当与输入分组对应的条目在其他表(OF)上存在时，基于其他表(OF)之前的查询结果确定命中了查询结果查阅条目122。

[0147] 例如，当与输入分组对应的条目在其他表(OF)的查询中存在时，设置标记“Z＝1”，并且当已经设置了标记“Z＝1”时，确定在TCAM(OF)的其他表(OF)查询结果查阅条目122中命中了条目。

[0148] 当确定其他表(OF)查询结果查阅条目122被命中时，开放流动作解析器107选择其他表(OF)的条目的动作作为针对该分组的动作。

[0149] (未命中条目)

[0150] 未命中条目123是当输入分组未命中任何TCAM(OF)条目时确定被命中的条目。

[0151] 这里，未命中条目123是具有任意模式的分组命中的条目。当输入分组未命中任何TCAM(OF)条目并仅命中未命中条目123时，开放流动作解析器107根据分组流的设置选择“分组输入”(向控制器询问分组的动作)或“通常”(使用传统功能部分的分组处理)作为针对分组的动作。

[0152] 当用户尝试改变每个表的动作优先级时，TCAM(OF)中条目的顺序可能改变。

[0153] (开放流表的控制的概述)

[0154] 图4是示出了通过根据本发明的控制器进行开放流表控制的示意图。控制器101、交换机102和协议控制部分103具有与其在图1中相同的机制和功能。表组113具有与其在图2中相同的机制和功能。

[0155] 交换机102包括TCAM 124、L2表125、L3表126和其他表127。

[0156] TCAM 124包括TCAM(OF)和TCAM(传统)。L2表125包括L2表(OF)和L2表(传统)。L3表126包括L3表(OF)和L3表(传统)。其他表127包括其他表(OF)和其他表(传统)。

[0157] TCAM(OF)、L2表(OF)、L3表(OF)和其他表(OF)构成开放流表。

[0158] TCAM(传统)、L2表(传统)、L3表(传统)和其他表(传统)构成传统表。

[0159] 通常地,交换机102上的TCAM 124、L2表125、L3表126和其他表127是物理上的一个表。

[0160] 根据本发明的交换机102具有将单个物理表(TCAM 124、L2表125、L3表126和其他表127)逻辑上划分为构成开放流表的表组113和构成传统表的表组128的功能。即,交换机102基于针对每个表定义的条件和处理内容,划分并逻辑上集成一个物理表(TCAM 124、L2表125、L3表126和其他表127),以构建开放流表(表组113)和传统表(表组128)。

[0161] 构成开放流表的表组113包括TCAM(OF)、L2表(OF)、L3表(OF)和其他表(OF)。

[0162] 构成传统表的表组128包括TCAM(传统)、L2表(传统)、L3表(传统)和其他表(传统)。

[0163] (开放流表的控制的概述)

[0164] 在下文中将描述通过图4中控制器进行开放流表控制的概述。

[0165] 控制器101可以通过协议控制部分103控制交换机102的开放流表。

[0166] 通常地,交换机102上的TCAM 124、L2表125、L3表126和其他表127是单个物理表。

[0167] 根据本发明的交换机102具有以下功能:剪裁并使用表资源的一部分用于开放流;以及逻辑上构建构成开放流表的表组113和构成传统表的表组128。即,交换机102基于TCAM 124、L2表125、L3表126和其他表127来构建逻辑上的开放流表(表组113)和传统表(表组128)。

[0168] 构成开放流表的表具有不同的可行的开放流功能。

[0169] 针对此原因,控制器101需要考虑如下内容来执行开放流表控制:“1:在每个表中可以执行什么开放流功能?”和“2:要控制的条目属于哪个构成开放流表的表?”。

[0170] 关于“1”,例如“控制器具有以下机制:先前输入(Input)每个表中可以实现的功能,并当尝试其他类型的控制时返回错误”,以及“交换机具有以下机制:当目标表不具有与控制相对应的功能时,从控制器返回针对控制命令的错误”。

[0171] 关于“2”,例如“将开放流表中特定范围的优先级(0-64k)分配给每个表,并且当控制器执行控制时,基于优先级范围确定要使用的表”,或“将ID分配给构成开放流表的每个表,当控制器执行控制时,基于ID确定要使用的表”。

[0172] (开放流表的控制的细节(1))

[0173] 图5是示出了根据本发明的控制器进行开放流表控制的第一方法的示例的细节的示意图。这里,将描述通过使用优先级范围来指定表的情况。

[0174] 控制器101、交换机102和开放流表管理部分106具有与其在图1中相同的机制和功能。表组113具有与其在图2中相同的机制和功能。

[0175] 开放流功能部分105的开放流表管理部分106将优先级范围分配给构成开放流表的表组113中的每个表。

[0176] 在基于优先级范围指定表的情况中，表的优先级范围必须不重叠。总的优先级范围必须不超过开放流中规定的优先级范围。

[0177] 控制器101基于分配给每个表的优先级范围中的值来指定表。

[0178] 开放流功能部分105的开放流表管理部分106基于控制器101指定的优先级范围中的值，确定要使用的表。

[0179] (开放流表的控制的细节(2))

[0180] 图6是示出了根据本发明的控制器进行开放流表控制的第二方法的示例的细节的示意图。这里，将描述通过使用表ID来指定表的情况。

[0181] 控制器101、交换机102和开放流表管理部分106具有与其在图1中相同的机制和功能。表组113具有与其在图2中相同的机制和功能。

[0182] 开放流功能部分105的开放流表管理部分106将表ID 129分配给构成开放流表的表组113中的每个表。

[0183] 在基于表ID来指定表的情况中，表ID必须不重叠。另一方面，针对表设置的优先级范围可能重叠。这是因为，根据表ID 129将每个表标识为单独的表，并且每个表的优先级是由开放流动作解析器107确定的。

[0184] 控制器101根据分配给每个表的表ID 129来指定表。

[0185] 开放流功能部分105的开放流表管理部分106基于由控制器101指定的表ID 129，确定要使用的表。

[0186] (分组进入时的处理操作示例(1))

[0187] 图7A和7B示出了根据本发明的分组进入时交换系统的第一操作示例。为了简洁，仅示出了与操作相关的功能块。步骤S101至S107示出了从分组进入时的整个操作的流。

[0188] (1)步骤S101

[0189] 控制器101向交换机102事先注册{匹配条件:目的IP＝AA，动作:丢掉}的TCAM(OF)条目(1)。交换机102将TCAM(OF)条目(1)注册到表组113的TCAM(OF)中。

[0190] 基于L1至L4的可选首部信息的组合，定义“匹配条件”。“动作”定义了例如对满足匹配条件的分组延迟/丢弃/重写首部信息的动作。

[0191] (2)步骤S102

[0192] 控制器101向交换机102事先注册{匹配条件:目的IP＝AA，动作:从端口1输出}的L3表(OF)条目(1)和{匹配条件:目的IP＝BB，动作:从端口2输出}的L3表(OF)条目(2)。交换机102将L3表(OF)条目(1)和L3表(OF)条目(2)注册到表组113的L3表(OF)中。

[0193] (3)步骤S103

[0194] 当分组进入时具有目的IP＝AA的分组流入开放流有效端口时，输入端口104向开放流功能部分105发送该分组。

[0195] (4)步骤S104

[0196] 当具有目的IP＝AA的分组流到开放流有效端口时，开放流功能部分105处理该分组。

[0197] (5)步骤S105

[0198] 首先，在开放流功能部分105中，L2/L3/其他表(OF)查询功能部分在L2/L3/其他表(OF)中查询进入分组。在此查询中，分组命中L3表(OF)条目(1)。

[0199] (6)步骤S106

[0200] L2/L3/其他表(OF)查询功能部分向开放流动作解析器107通知查询结果。在图7B中，开放流动作解析器107和TCAM(OF)查询功能部分116作为一个功能块(TCAM(OF)查询&开放流动作解析器)而示出。

[0201] (7)步骤S107

[0202] TCAM(OF)查询功能部分116在TCAM(OF)中查询进入分组。在此查询中，分组首先命中TCAM(OF)条目(1)。此时，TCAM(OF)的查询处理结束。TCAM(OF)查询功能部分116向开放流动作解析器107通知查询结果。

[0203] (8)步骤S108

[0204] 开放流动作解析器107从L2/L3/其他表(OF)查询功能部分和TCAM(OF)查询功能部分116中的每一个接收查询结果，并根据表间优先级确定针对进入分组的动作。这里，TCAM(OF)的优先级是最高的。针对此原因，开放流动作解析器107确定将TCAM(OF)条目(1)的动作(丢掉)作为针对进入分组的动作，并向动作功能部分111通知所确定的动作(丢掉)。

[0205] (9)步骤S109

[0206] 动作功能部分111执行由开放流动作解析器107确定的动作(丢掉)。这里，因为动作是“丢掉”，动作功能部分111不输出该分组。动作功能部分111丢弃进入分组和属于相同流的随后的分组。

[0207] (分组进入时的处理操作示例(2))

[0208] 图8A和8B示出了根据本发明的分组进入时交换系统的第二操作示例。为了简洁，仅示出了与操作相关的功能块。步骤S201至S210示出了从分组进入时的整个操作的流。

[0209] (1)步骤S201

[0210] 控制器101向交换机102事先注册{匹配条件：目的IP＝AA，动作：丢掉}的TCAM(OF)条目(1)。交换机102将TCAM(OF)条目(1)注册到表组113的TCAM(OF)中。

[0211] (2)步骤S202

[0212] 控制器101向交换机102事先注册{匹配条件：目的IP＝AA，动作：从端口1输出}的L3表(OF)条目(1)和{匹配条件：目的IP＝BB，动作：从端口2输出}的L3表(OF)条目(2)。交换机102将L3表(OF)条目(1)和L3表(OF)条目(2)注册到表组113的L3表(OF)中。

[0213] (3)步骤S203

[0214] 当分组进入时具有目的IP＝BB的分组流入开放流有效端口时，输入端口104向开放流功能部分105发送该分组。

[0215] (4)步骤S204

[0216] 当具有目的IP＝BB的分组流到开放流有效端口时，开放流功能部分105处理该分组。

[0217] (5)步骤S205

[0218] 首先，在开放流功能部分105中，L2/L3/其他表(OF)查询功能部分在L2/L3/其他表(OF)中查询分组。在此查询中，分组命中L3表(OF)条目(2)。

[0219] (6)步骤S206

[0220] L2/L3/其他表(OF)查询功能部分向开放流动作解析器107通知查询结果。在图8B中,开放流动作解析器107和TCAM(OF)查询功能部分116作为一个功能块(TCAM(OF)查询&开放流动作解析器)而示出。

[0221] (7)步骤S207

[0222] TCAM(OF)查询功能部分116在TCAM(OF)中查询进入分组。然而,因为进入分组是具有目的IP＝BB的分组,该分组未命中TCAM(OF)中存在的任何条目。此时,TCAM(OF)的查询处理结束。TCAM(OF)查询功能部分116向开放流动作解析器107通知查询结果。此时,TCAM(OF)的查询处理结束。TCAM(OF)查询功能部分116向开放流动作解析器107通知查询结果。

[0223] (8)步骤S208

[0224] 开放流动作解析器107从L2/L3/其他表(OF)查询功能部分和TCAM(OF)查询功能部分116中的每一个接收查询结果,并根据表间优先级确定进入分组的动作。这里,仅在L3表(OF)条目(2)中存在目标条目。因此,开放流动作解析器107将L3表(OF)条目(2)的动作(从端口2输出)确定为针对进入分组的动作,并向动作功能部分111通知所确定的动作(从端口2输出)。

[0225] (9)步骤S209

[0226] 动作功能部分111执行由开放流动作解析器107确定的动作(从端口2输出)。

[0227] (10)步骤S210

[0228] 动作功能部分111向具有端口2的输出端口输出进入分组和属于相同流的随后的分组。

[0229] (11)步骤S211

[0230] 输出端口向端口2输出从动作功能部分111输出的分组。

[0231] (12)步骤S212

[0232] 从端口2输出的分组向网络流出,并向目的IP＝BB发送。

[0233] 在图8A和8B中,当进入分组命中多个表中的任意条目时,L2/L3/其他表(OF)查询功能部分向开放流动作解析器107通知所有命中条目作为查询结果。当没有进入分组命中最高优先级的TCAM(OF)中的任意条目时,开放流动作解析器107在L2/L3/其他表(OF)查询功能部分的查询结果中采用较高优先级的表中命中的条目。

[0234] 在图8A和8B中,当进入分组命中与多个表中的条目时,L2/L3/其他表(OF)查询功能部分可以根据表间优先级117向开放流动作解析器107通知在这些表中的较高优先级的表中命中的条目,作为查询结果。

[0235] (开放流表控制示例(1))

[0236] 图9示出了在基于优先级范围指定表的情况下,控制器进行开放流表控制的第一方法的示例。

[0237] 注册信息132是从控制器101到交换机102的开放流表管理部分106的信息。注册结果是构成开放流表的表组113的注册结果。

[0238] 当构成开放流表的表组113被假定为一个表时,交换机102根据优先级控制期望的表。

[0239] (1)步骤S301

[0240] 控制器101向交换机102注册{优先级:50001,匹配条件:XXXX,动作:YYYY}的条目。

[0241] (2)步骤S302

[0242] 当从控制器101注册条目时，交换机102根据条目的优先级选择L2表(OF)作为表，并将该条目注册到L2表(OF)中。

[0243] (开放流表控制示例(2))

[0244] 图10示出了在基于表ID指定表的情况下，控制器进行开放流表控制的第二方法的示例。

[0245] 注册信息是从控制器到开放流表的注册信息。注册结果135是构成开放流表的表组的注册结果。

[0246] 当构成开放流表的表组中的表被假定为不同的开放流表时，交换机102根据表ID控制期望的表。

[0247] (1)步骤S401

[0248] 控制器101向交换机102注册{表ID:#2，优先级:1，匹配条件:XXXX，动作:YYYY}的条目。

[0249] (2)步骤S402

[0250] 当从控制器101注册条目时，交换机102根据表ID选择L2表(OF)作为表，并将该条目注册到L2表(OF)中。

[0251] (第一示例性实施例的特征)

[0252] 在第一示例性实施例中，开放流动作解析器可以集成多个表的资源，并相互比较表的优先级以解析动作。

[0253] 因此，可以通过使用交换机的多个表的资源构建大容量的开放流表。从而，交换机可以控制大量流。

[0254] 在此示例性实施例中，可以基于在使用开放流表作为一个大容量的开放流表的情况中的“优先级范围”或在使用构成开放流表的表组中的每个表作为多个不同的开放流表的情况中的“表ID”，来标识构成开放流表的每个表。

[0255] 因此，由交换机的多个表构成的开放流表可以用作一个大容量的开放流表或多个不同的开放流表。从而，可以灵活控制从多个表配置的开放流表。

[0256] [第二示例性实施例]

[0257] 在下文中将参考附图描述本发明的第二示例性实施例。第二示例性实施例是当TCAM(OF)查询功能部分不包括开放流动作解析器时开放流动作解析器的示例性实施例。

[0258] <示例1>

[0259] 图11是示出了当通过“TCAM(OF)查询功能部分116”和“L2/L3/其他表(OF)查询功能部分115”并按此顺序执行查询时，开放流功能部分105的细节。

[0260] 开放流功能部分105、开放流表管理部分106、开放流动作解析器107、动作功能部分111、表组113、查询功能部分114、L2/L3/其他表(OF)查询功能部分115和TCAM(OF)查询功能部分116具有与其在图2中相同的机制和功能。

[0261] 参考图11，将描述当通过“TCAM(OF)查询功能部分116”和“L2/L3/其他表(OF)查询功能部分115”并按此顺序执行查询时，开放流功能部分105的操作。

[0262] 图11示出了以下情况：在开放流表管理部分106的查询功能部分114中，TCAM(OF)查询功能部分116最终不执行查询。

[0263] 在图11中，因为通过“TCAM(OF)查询功能部分116”和“L2/L3/其他表(OF)查询功能部分115”并按此顺序执行查询，则TCAM(OF)查询功能部分116可以不包括开放流动作解析器107。

[0264] <示例2>

[0265] 图12是示出了当L2/L3/其他表(OF)查询功能部分115和TCAM(OF)查询功能部分116同时执行查询时，开放流功能部分105的细节的示意图。

[0266] 开放流功能部分105、开放流表管理部分106、开放流动作解析器107、动作功能部分111、表组113、查询功能部分114、L2/L3/其他表(OF)查询功能部分115和TCAM(OF)查询功能部分116具有与其在图2中相同的机制和功能。

[0267] 参考图12，将描述当L2/L3/其他表(OF)查询功能部分115和TCAM(OF)查询功能部分116同时执行查询时，开放流功能部分105的操作。

[0268] 图12示出了以下情况：在开放流表管理部分106的查询功能部分114中，TCAM(OF)查询功能部分116最终不执行查询。

[0269] 因为L2/L3/其他表(OF)查询功能部分115和TCAM(OF)查询功能部分116同时执行查询，则TCAM(OF)查询功能部分116可以不包括开放流动作解析器107。

[0270] <示例3>

[0271] 图13示出了当TCAM(OF)查询功能部分可以不包括开放流动作解析器时开放流动作解析器107的示例性实施例。

[0272] 开放流动作解析器107和动作功能部分111与其在图2中具有相同的机制和功能。

[0273] 开放流动作解析器107包括查询接收部分(查询接收器)130和处理解析部分(动作解析器)131。

[0274] 查询接收部分(查询接收器)130接收每个表的查询结果。处理解析部分(动作解析器)131基于每个表的查询结果，确定开放流处理。

[0275] 参考图13，将描述当TCAM(OF)查询功能部分可以不包括开放流动作解析器107时开放流动作解析器107的操作。

[0276] 当TCAM(OF)查询功能部分可以不包括开放流动作解析器107时，图13中示出的开放流动作解析器107安装在交换机102中。

[0277] 在开放流动作解析器107中，查询接收部分(查询接收器)130接收每个表的查询结果。处理解析部分(动作解析器)131基于每个表的查询结果和先前设置的表间优先级，确定开放流的动作，并向动作功能部分111通知所确定的动作。

[0278] [其他示例性实施例]

[0279] 虽然在每个上述示例性实施例中，将交换机102描述为开放流交换机，开放流交换机仅是示例。实际上，交换机不限于开放流交换机，并且本发明也可以适用于具有与开放流交换机相同机制和功能的任何交换机。

[0280] <本发明的特征>

[0281] 本发明涉及通过集成多个表来扩展开放流表的方法。

[0282] 根据本发明，在由交换机的单个表(主要是TCAM)构成的开放流表中，通过从交换机的多个表构建开放流表来实现开放流表中的流条目的数量的扩展。

[0283] 在开放流技术中，传送功能和控制功能已经被安装在相同的NW设备(路由器/交换

机等)中并且相互独立,传送功能仍然保留在NW设备中,而控制功能由外部控制器代替。控制器根据开放流协议,远程地操作NW设备中的开放流表以控制NW设备的行为。开放流表由包括三类信息({匹配条件、动作、统计信息})的流条目组构成。在开放流技术中,匹配条件定义了要控制的流,而动作和统计信息可以在流的单元中获取。

[0284] 以下作出开放流的{匹配条件、动作、统计信息}的概述。

[0285] (匹配条件)

[0286] “进入端口(输入端口)”/“源MAC(源MAC地址)”/“目的MAC(目的地MAC地址)”/“以太类型(类型/领域)”/“VLAN ID(虚拟LAN标识信息)”/“VLAN优先级(虚拟LAN优先级)”/“源IP(源IP地址)”/“目的IP(目的地IP地址)”/“IP协议(IP协议号)”/“IP ToS(上6位)”/源端口(源端口号)”/“目的端口(目的地端口号)”。

[0287] (动作)

[0288] “转发(从物理端口输出)”/“所有(从除了输出端口以外的所有端口输出)”/“控制器(向控制器输出)”/“本地(向设备的本地栈输出)”/“表(根据开放流表中的内容输出)”/“输入端口(从输入端口输出)”/“通常(使用传统表中的内容输出)”/“泛滥(从除了输入端口和展开树的阻塞端口以外的所有端口输出)”/“丢掉(丢弃分组)”/“修改字段(重写分组的首部信息)”。

[0289] 例如,在“修改字段”,“VLAN ID”、“VLAN优先级(优先级)”、“源MAC”、“目的MAC”、“源IP”、“目的IP”、“IP ToS”、“源端口”和“目的端口”的情况下可以重写。

[0290] (统计信息)

[0291] “表”、“流”、“物理端口”和“队列(队列)”单元中各种类型的统计信息。

[0292] 根据本发明,通过使用交换机的多个表构成开放流表,可以实现如设备的开放流表中的流条目数量的扩展,而不增加交换机的表(主要是TCAM)自身的容量。即,交换机中多个表可以用作控制器侧的大容量的开放流表。

[0293] 具体地,通过吸收每个表的功能之间的差异来实现向开放流表的集成。因为交换机的多个表中每一个具有原始用途(例如L2表中的L2中继、L3表中的L3中继),所有资源未被使用,并且部分资源被切掉并使用。

[0294] 如上所述,本发明的特征是“吸收多个表的匹配条件/动作差异并将表集成到开放流表中”,以及“提供特定动作的确定方法”。

[0295] 根据本发明,交换机的多个表资源的一部分用作开放流表。

[0296] 根据本发明,根据可行功能(匹配条件/动作),将交换机的每个表看作“功能受限的开放流表资源”。

[0297] 根据本发明,开放流动作解析器吸收每个表资源的功能(匹配条件/动作)之间的差异,并将资源集成作为开放流表资源。

[0298] 根据本发明,开放流动作解析器基于包括TCAM的表的优先级(优先级)来确定动作。

[0299] 在其中最终执行TCAM(OF)查询的交换机中,开放流动作解析器包括在TCAM(OF)查询功能部分中。

[0300] 根据本发明,配置有多个表的开放流表由控制器灵活地控制。

[0301] 根据本发明,当多个表用作一个开放流表时,基于优先级范围来标识表。

[0302] 根据本发明，当多个表用作开放流表时，基于表ID来标识表。

[0303] 虽然已经详细描述了本发明的示例性实施例，本发明不限于上述示例性实施例，并且不背离本发明的主题的修改落入本发明的范围中。

[0304] 本申请要求基于日本专利申请号JP 2010-200690的优先权。其公开以引用方式并入本文。

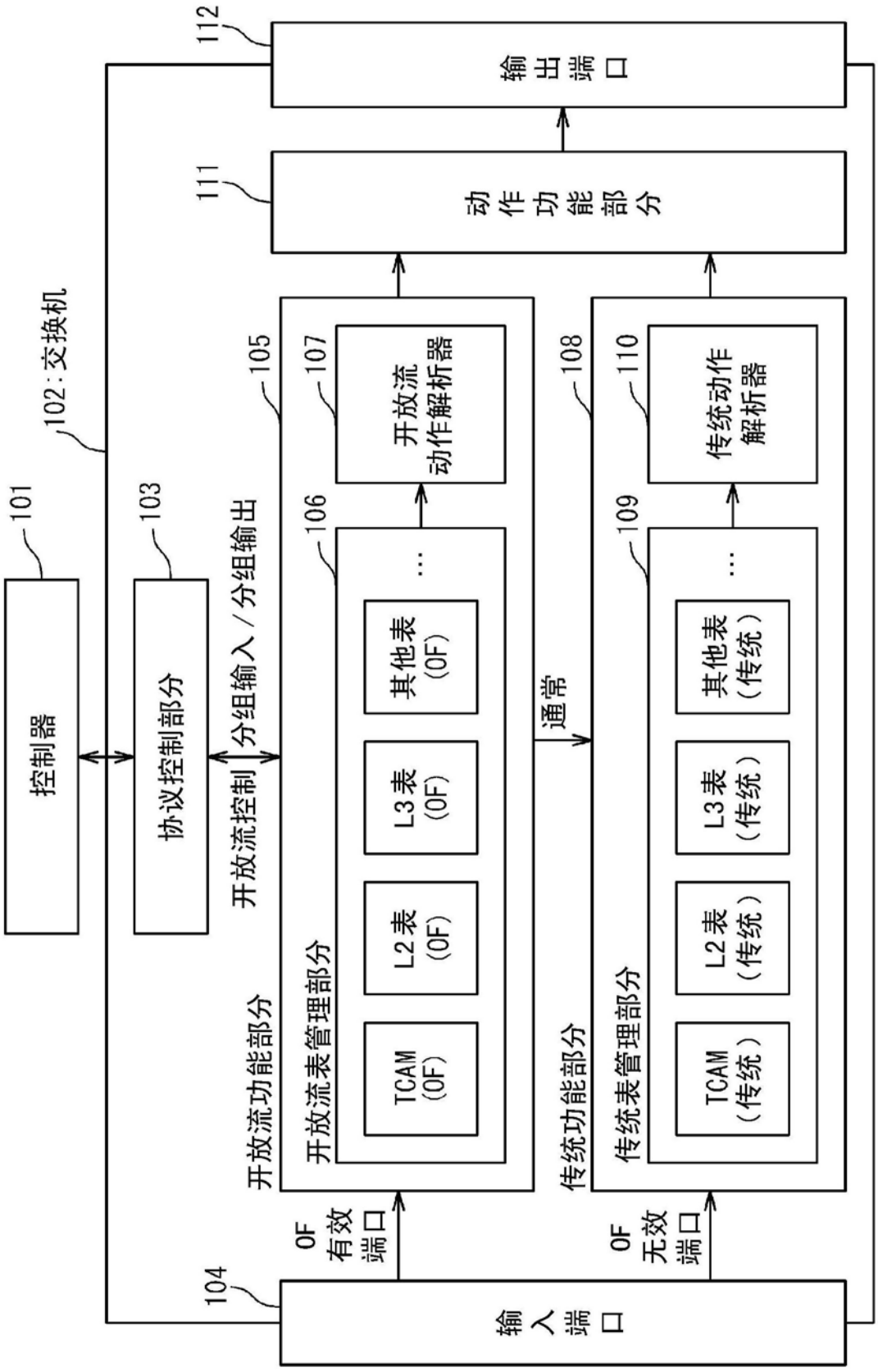

112
输出端口
111
动作功能部分
102:交换机
105
107
开放流动作解析器
108
110
传统动作解析器
101
103
106
109
控制器
协议控制部分
开放流控制
分组输入/分组输出
其他表(OF)
L3表(OF)
L2表(OF)
TCAM(OF)
通常
其他表(传统)
L3表(传统)
L2表(传统)
TCAM(传统)
开放流功能部分
开放流表管理部分
传统功能部分
传统表管理部分
OF有效端口
OF无效端口
104
输入端口

图1

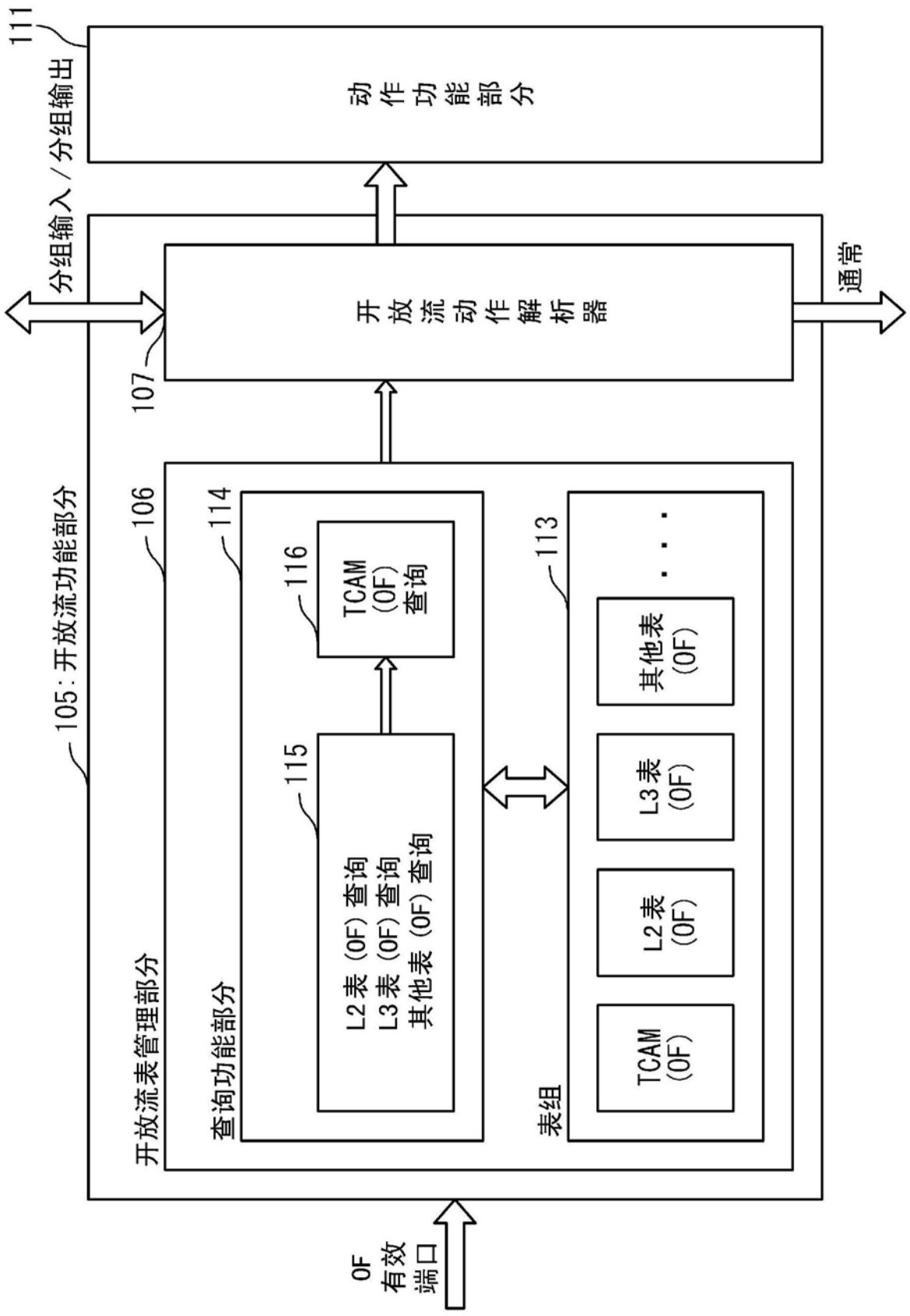

111
动作功能部分
分组输入/分组输出
通常
开放流动作解析器
107
105:开放流功能部分
106
开放流表管理部分
114
查询功能部分
116
TCAM (OF) 查询
115
L2表 (OF) 查询
L3表 (OF) 查询
其他表 (OF) 查询
113
表组
TCAM (OF)
L2表 (OF)
L3表 (OF)
其他表 (OF)
. . .
OF 有效端口

图2

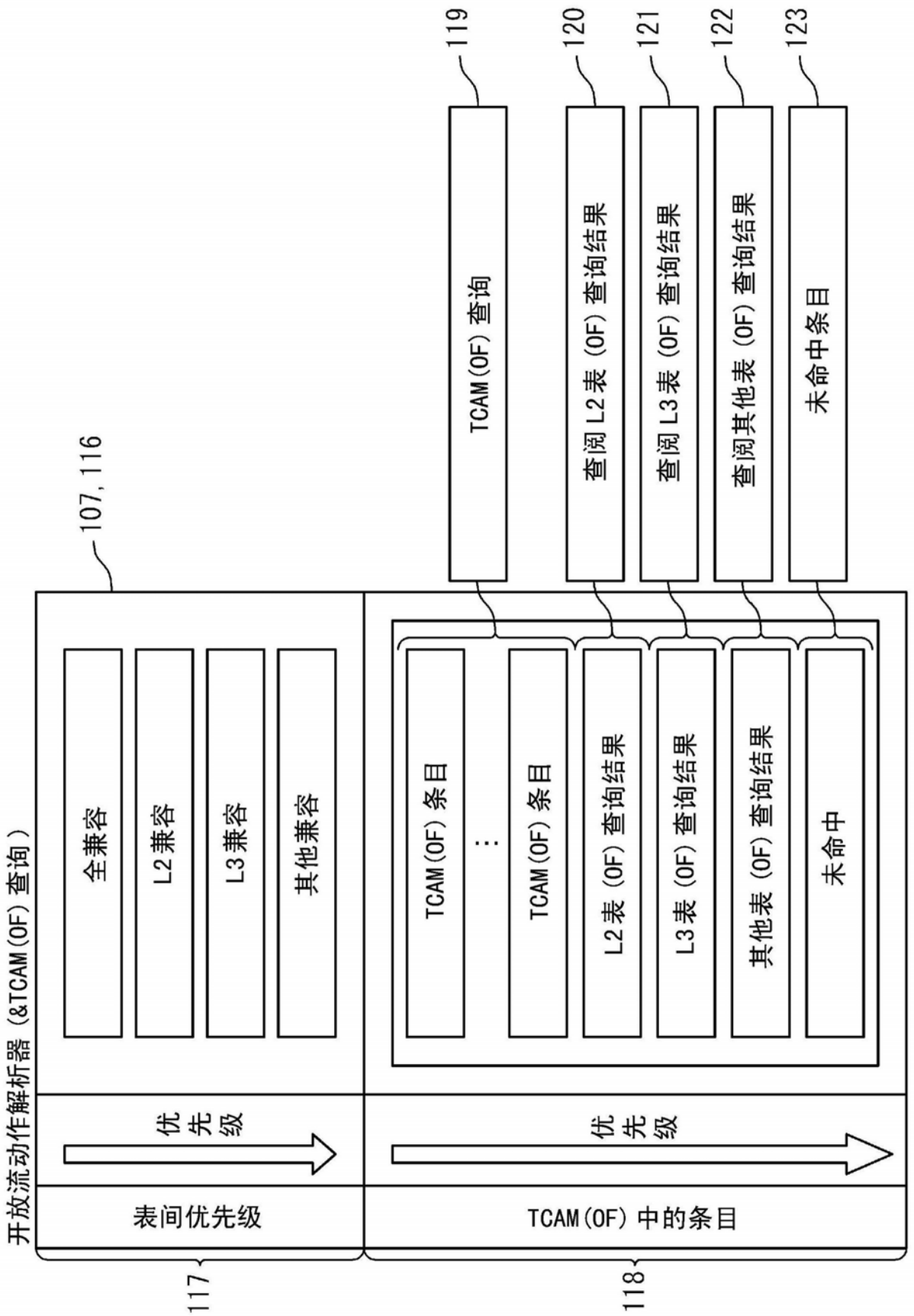

开放流动作解析器（&TCAM(OF)查询）
107,116
全兼容
L2兼容
L3兼容
其他兼容
优先级
表间优先级
117
TCAM(OF)查询
119
查阅L2表（OF）查询结果
120
查阅L3表（OF）查询结果
121
查阅其他表（OF）查询结果
122
未命中条目
123
TCAM(OF)条目
TCAM(OF)条目
L2表（OF）查询结果
L3表（OF）查询结果
其他表（OF）查询结果
未命中
优先级
TCAM(OF)中的条目
118

图3

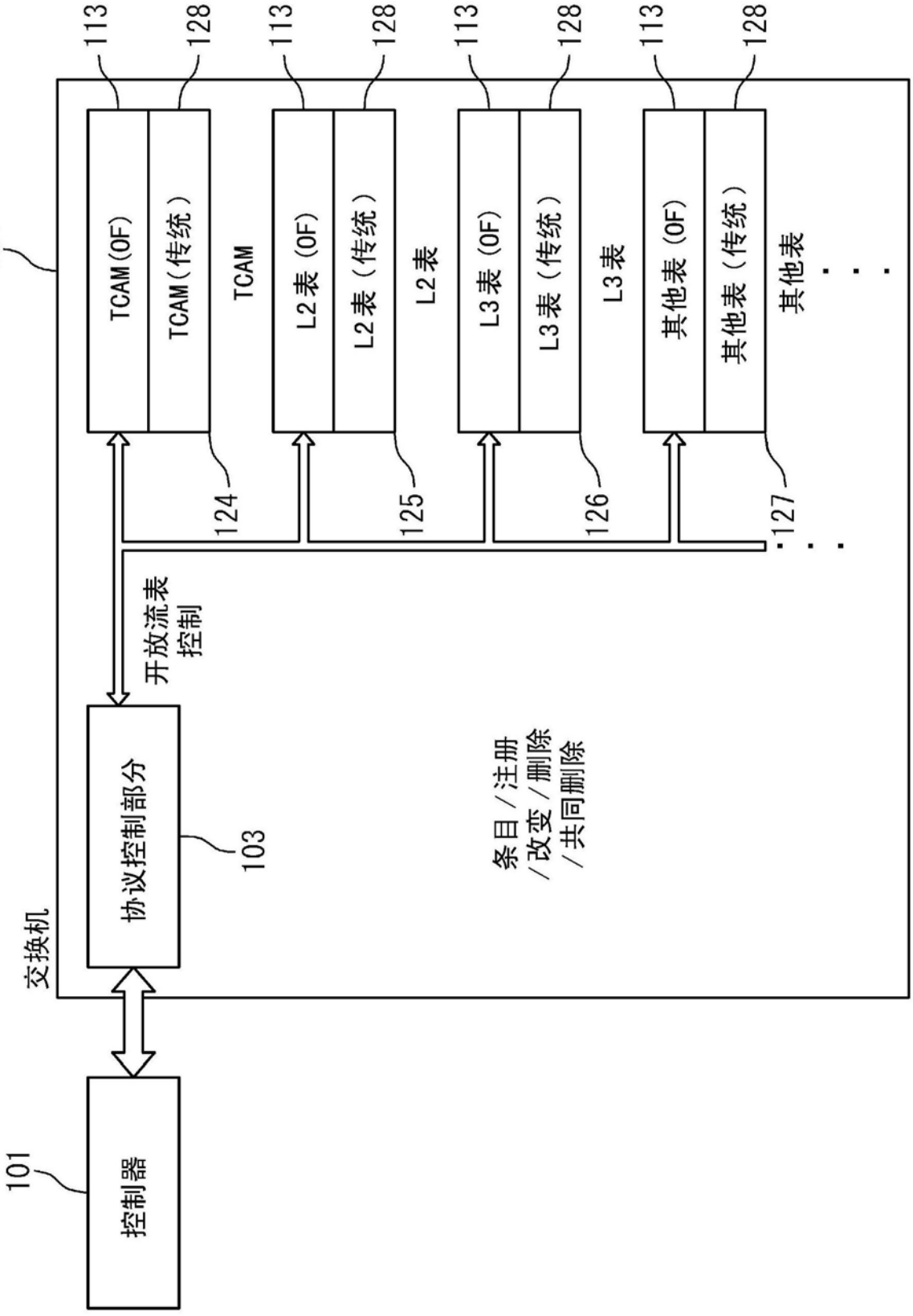

101
控制器
交换机
102
协议控制部分
103
开放流表控制
条目/注册/改变/删除/共同删除
TCAM(OF)
TCAM(传统)
TCAM
L2表(OF)
L2表(传统)
L2表
L3表(OF)
L3表(传统)
L3表
其他表(OF)
其他表(传统)
其他表
113
128
124
125
126
127

图4

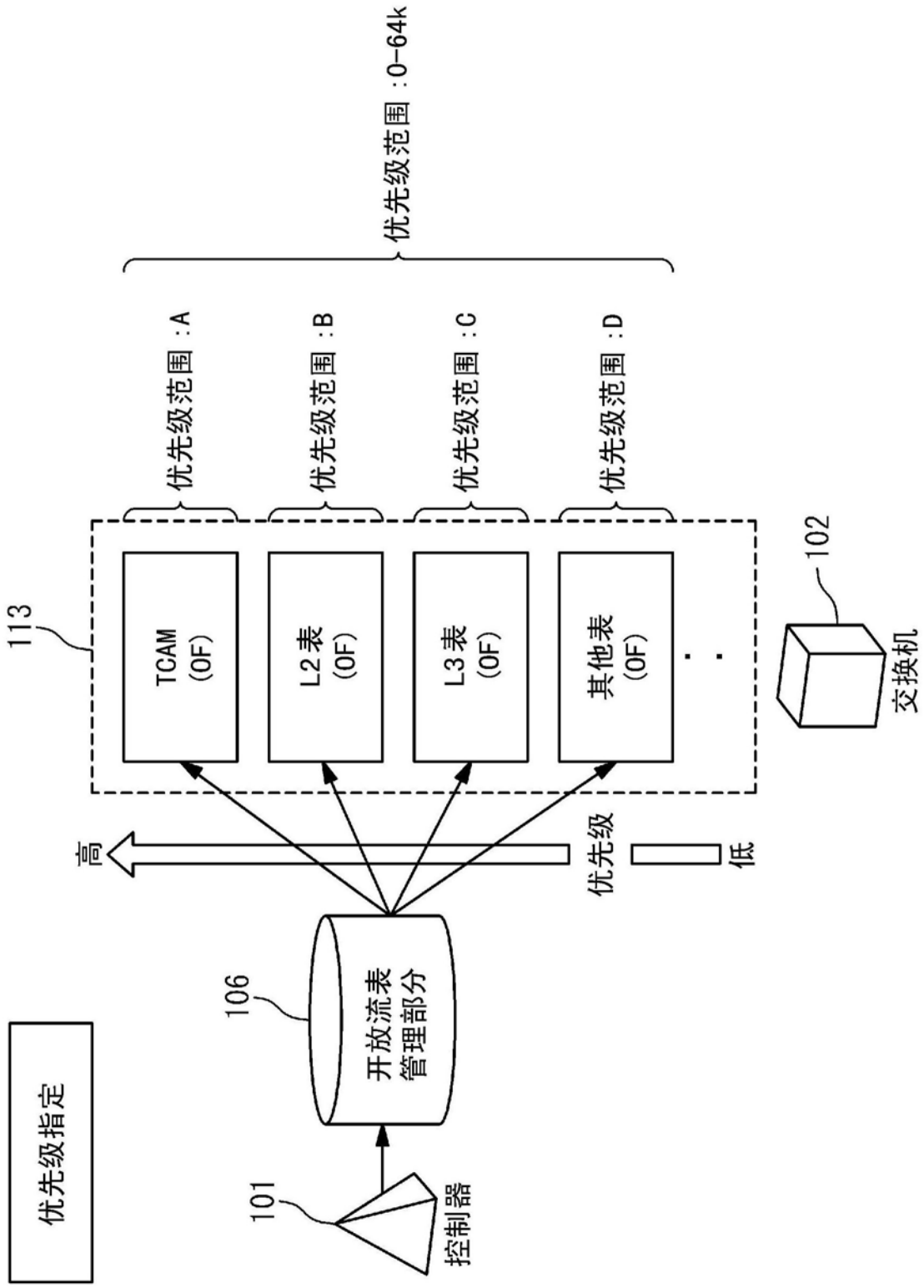

优先级范围 :0-64k
优先级范围 :A
优先级范围 :B
优先级范围 :C
优先级范围 :D
113
TCAM (OF)
L2表 (OF)
L3表 (OF)
其他表 (OF)
102
交换机
高
优先级
低
106
开放流表管理部分
优先级指定
101
控制器

图5

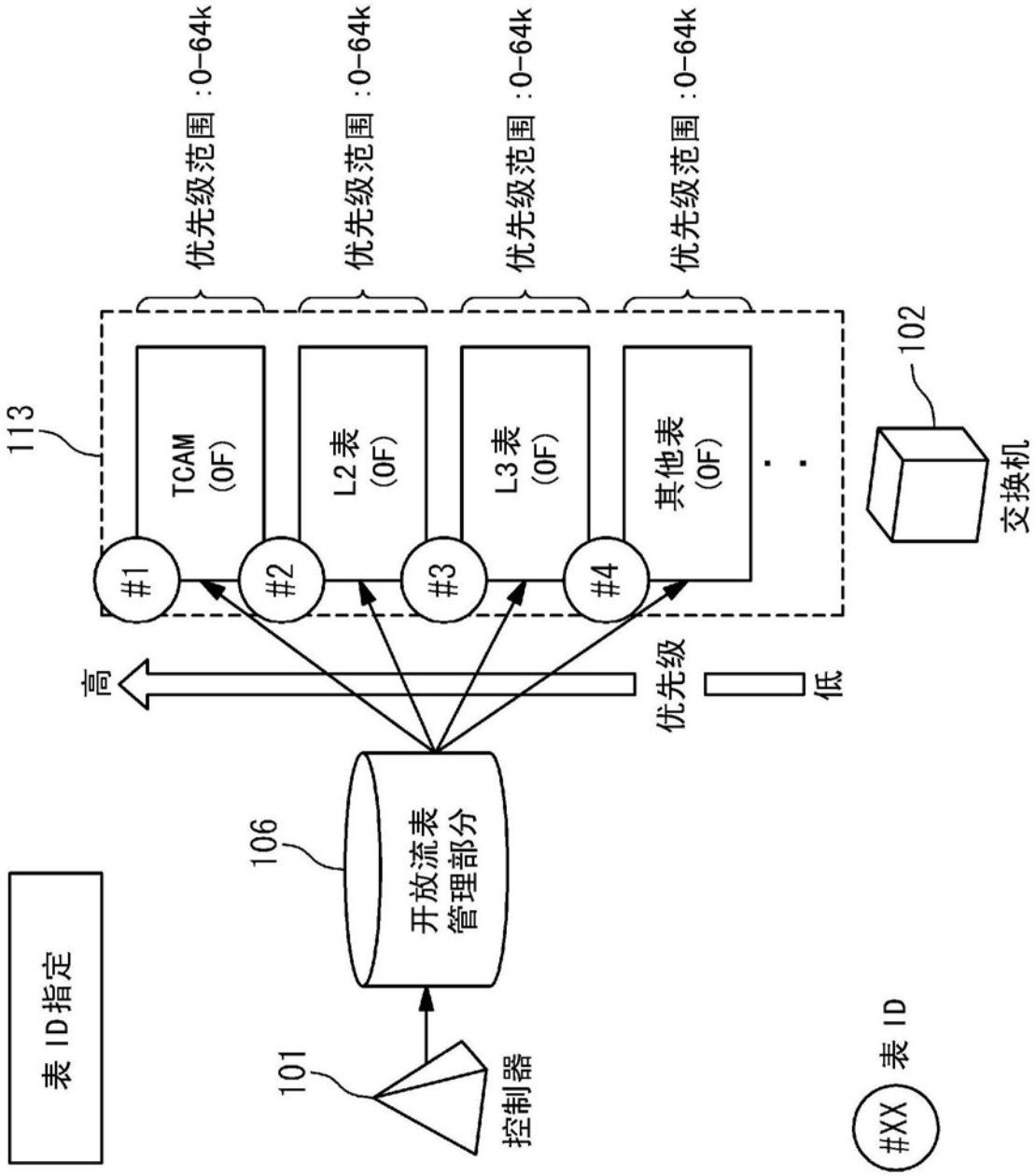

优先级范围 :0-64k
优先级范围 :0-64k
优先级范围 :0-64k
优先级范围 :0-64k
113
TCAM (OF)
L2表 (OF)
L3表 (OF)
其他表 (OF)
#1
#2
#3
#4
102
交换机
高
优先级
低
106
开放流表管理部分
表 ID指定
101
控制器
#XX 表 ID

图6

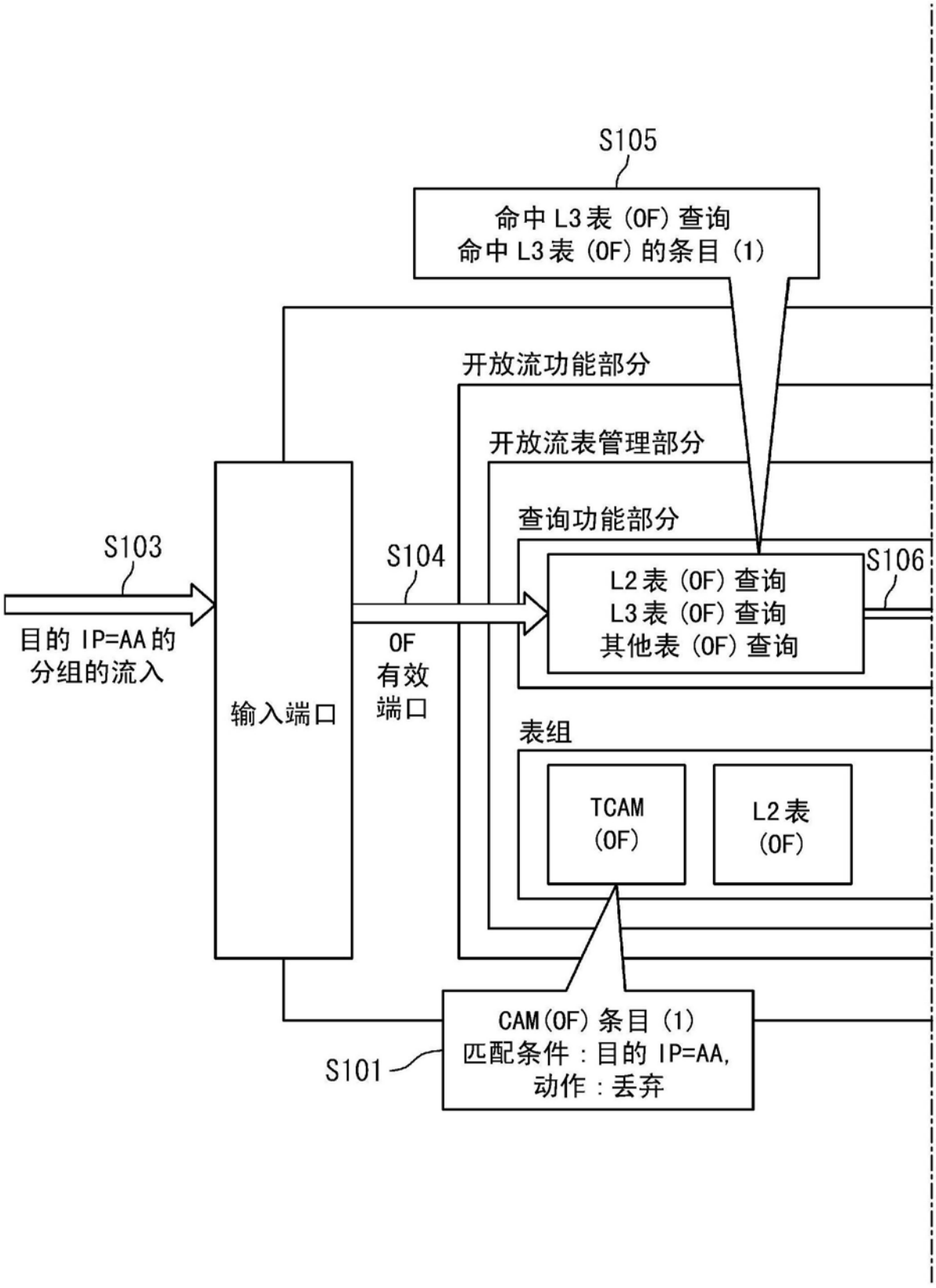

S105
命中 L3 表 (OF) 查询
命中 L3 表 (OF) 的条目 (1)
开放流功能部分
开放流表管理部分
查询功能部分
S103
S104
S106
L2 表 (OF) 查询
L3 表 (OF) 查询
其他表 (OF) 查询
目的 IP=AA 的
分组的流入
输入端口
OF
有效
端口
表组
TCAM
(OF)
L2 表
(OF)
CAM(OF) 条目 (1)
匹配条件：目的 IP=AA,
动作：丢弃
S101

图7A

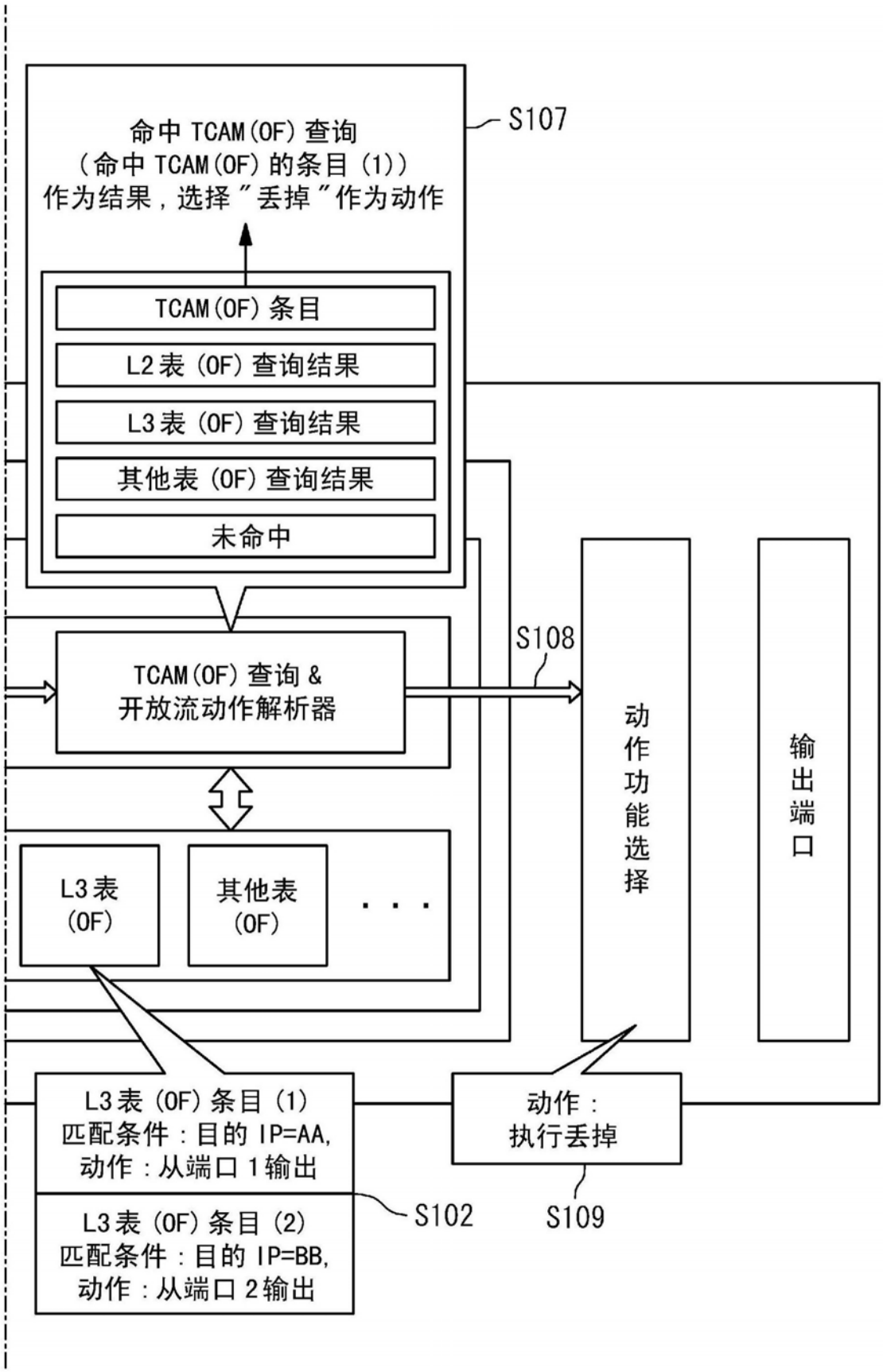

命中 TCAM(OF) 查询
(命中 TCAM(OF) 的条目 (1))
作为结果，选择"丢掉"作为动作
S107
TCAM(OF) 条目
L2 表 (OF) 查询结果
L3 表 (OF) 查询结果
其他表 (OF) 查询结果
未命中
TCAM(OF) 查询 &
开放流动作解析器
S108
动作功能选择
输出端口
L3 表
(OF)
其他表
(OF)
. . .
L3 表 (OF) 条目 (1)
匹配条件：目的 IP=AA,
动作：从端口 1 输出
L3 表 (OF) 条目 (2)
匹配条件：目的 IP=BB,
动作：从端口 2 输出
S102
动作：
执行丢掉
S109

图7B

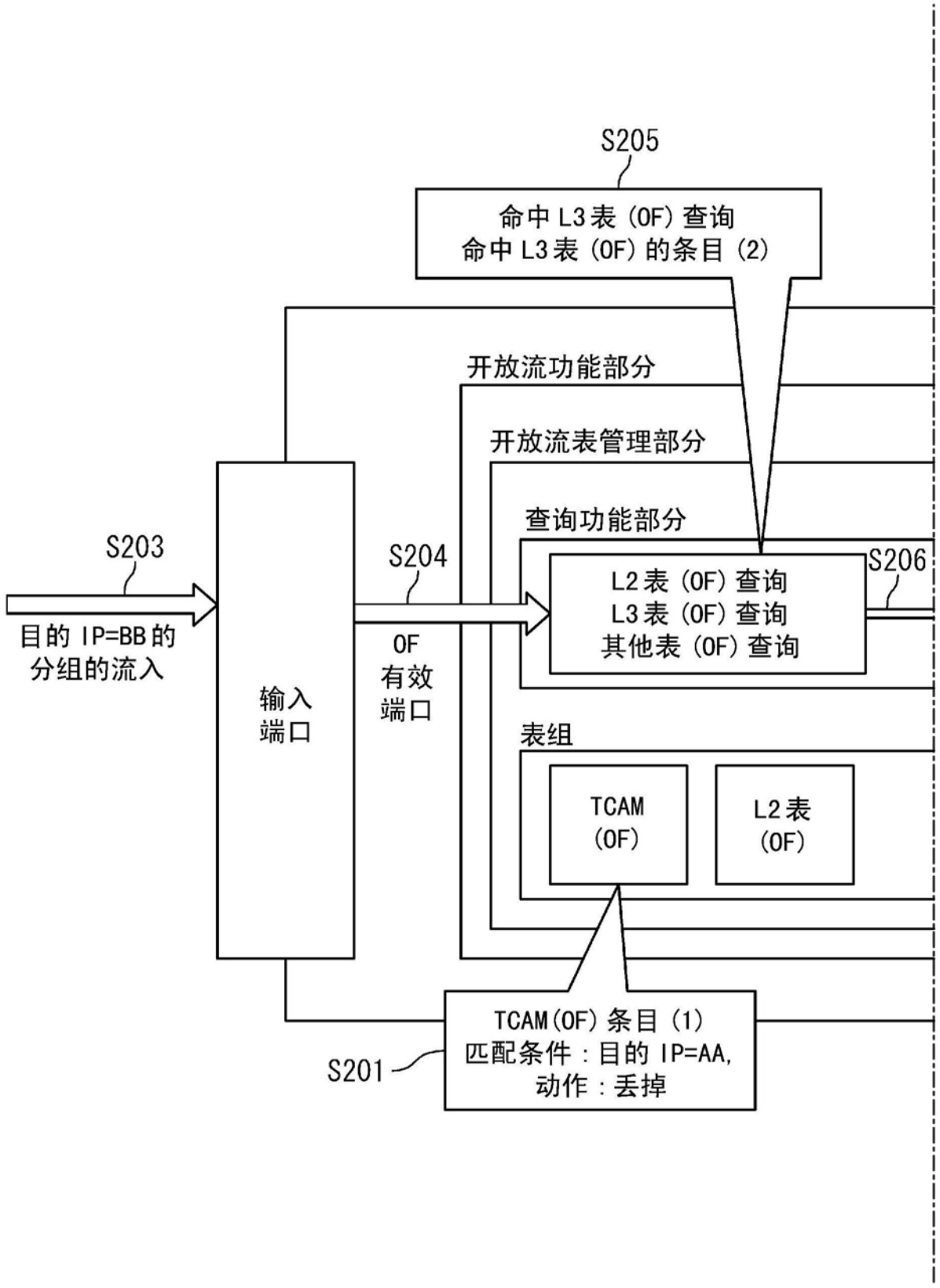

S205
命中 L3 表 (OF) 查询
命中 L3 表 (OF) 的条目 (2)
开放流功能部分
开放流表管理部分
查询功能部分
S203
S204
S206
目的 IP=BB 的
分组的流入
输入
端口
OF
有效
端口
L2 表 (OF) 查询
L3 表 (OF) 查询
其他表 (OF) 查询
表组
TCAM
(OF)
L2 表
(OF)
TCAM(OF) 条目 (1)
匹配条件：目的 IP=AA,
动作：丢掉
S201

图8A

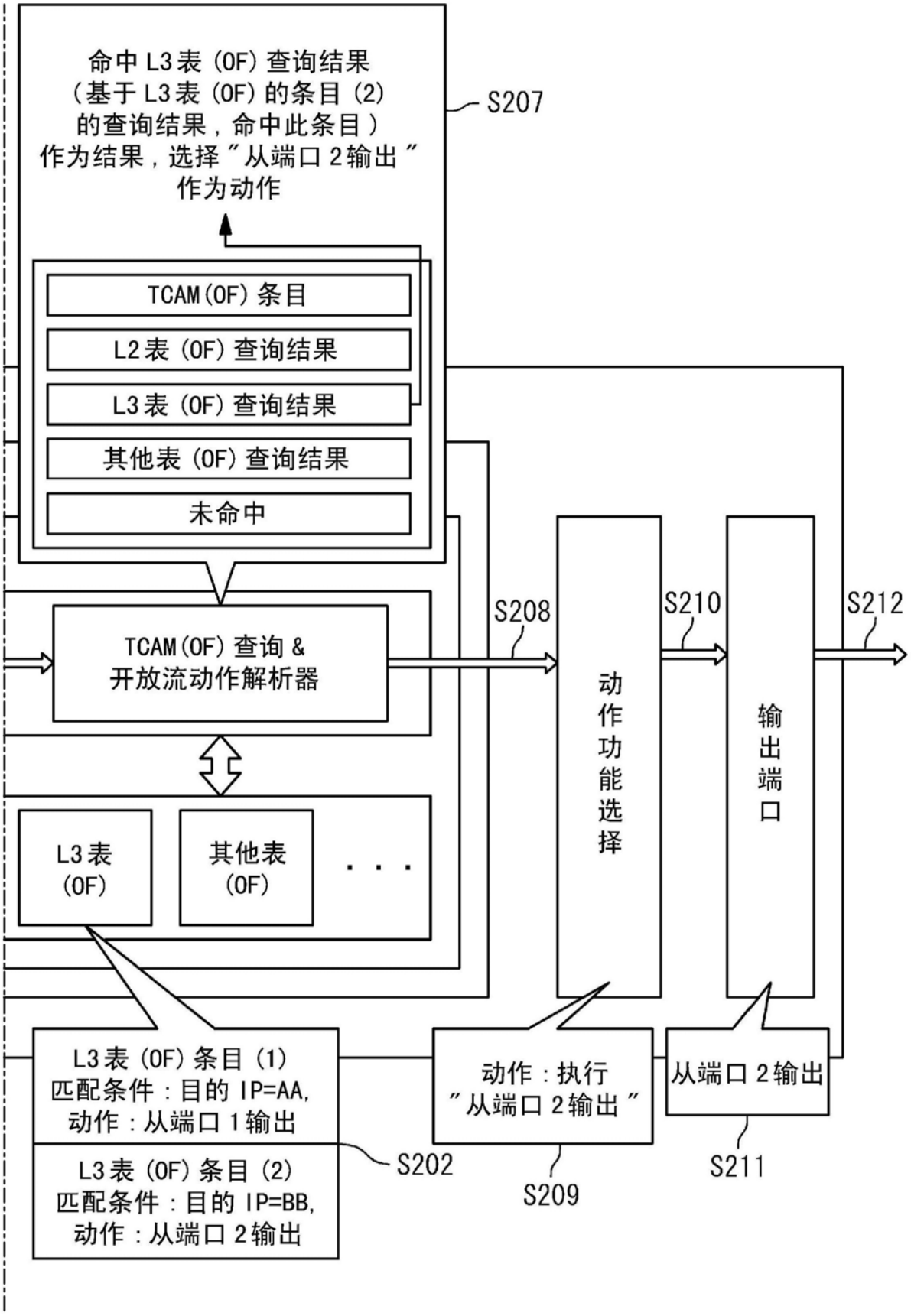

命中 L3 表 (OF) 查询结果
(基于 L3 表 (OF) 的条目 (2)
的查询结果，命中此条目)
作为结果，选择"从端口 2 输出"
作为动作
S207
TCAM(OF) 条目
L2 表 (OF) 查询结果
L3 表 (OF) 查询结果
其他表 (OF) 查询结果
未命中
TCAM(OF) 查询 &
开放流动作解析器
S208
S210
S212
动作功能选择
输出端口
L3 表
(OF)
其他表
(OF)
. . .
L3 表 (OF) 条目 (1)
匹配条件：目的 IP=AA,
动作：从端口 1 输出
L3 表 (OF) 条目 (2)
匹配条件：目的 IP=BB,
动作：从端口 2 输出
S202
动作：执行
"从端口 2 输出"
S209
从端口 2 输出
S211

图8B

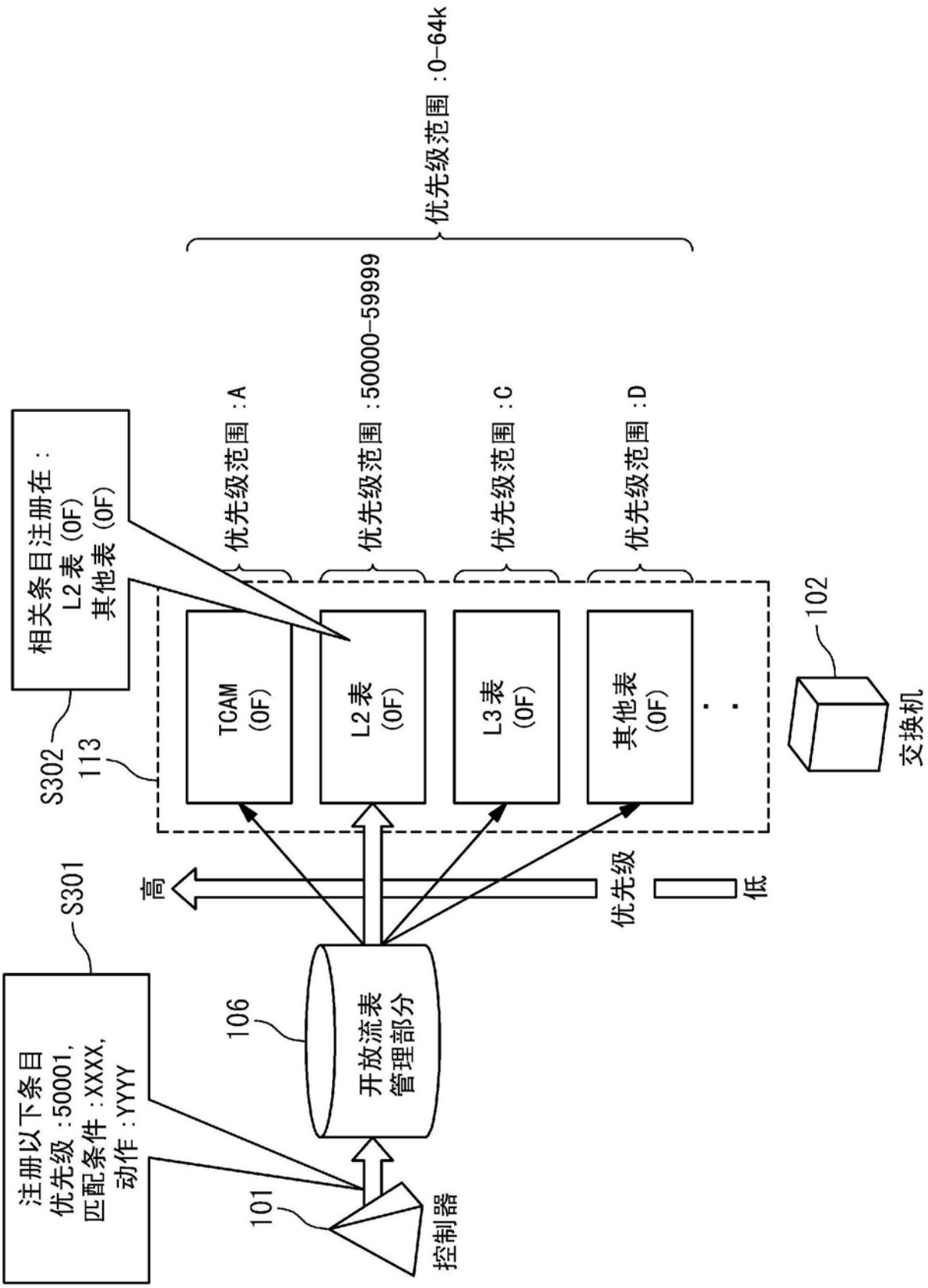

优先级范围 :0-64k
优先级范围 :A
优先级范围 :50000-59999
优先级范围 :C
优先级范围 :D
相关条目注册在：
L2表 (OF)
其他表 (OF)
S302
113
TCAM
(OF)
L2表
(OF)
L3表
(OF)
其他表
(OF)
102
交换机
高
优先级
低
S301
注册以下条目
优先级 :50001,
匹配条件 :XXXX,
动作 :YYYY
106
开放流表
管理部分
101
控制器

图9

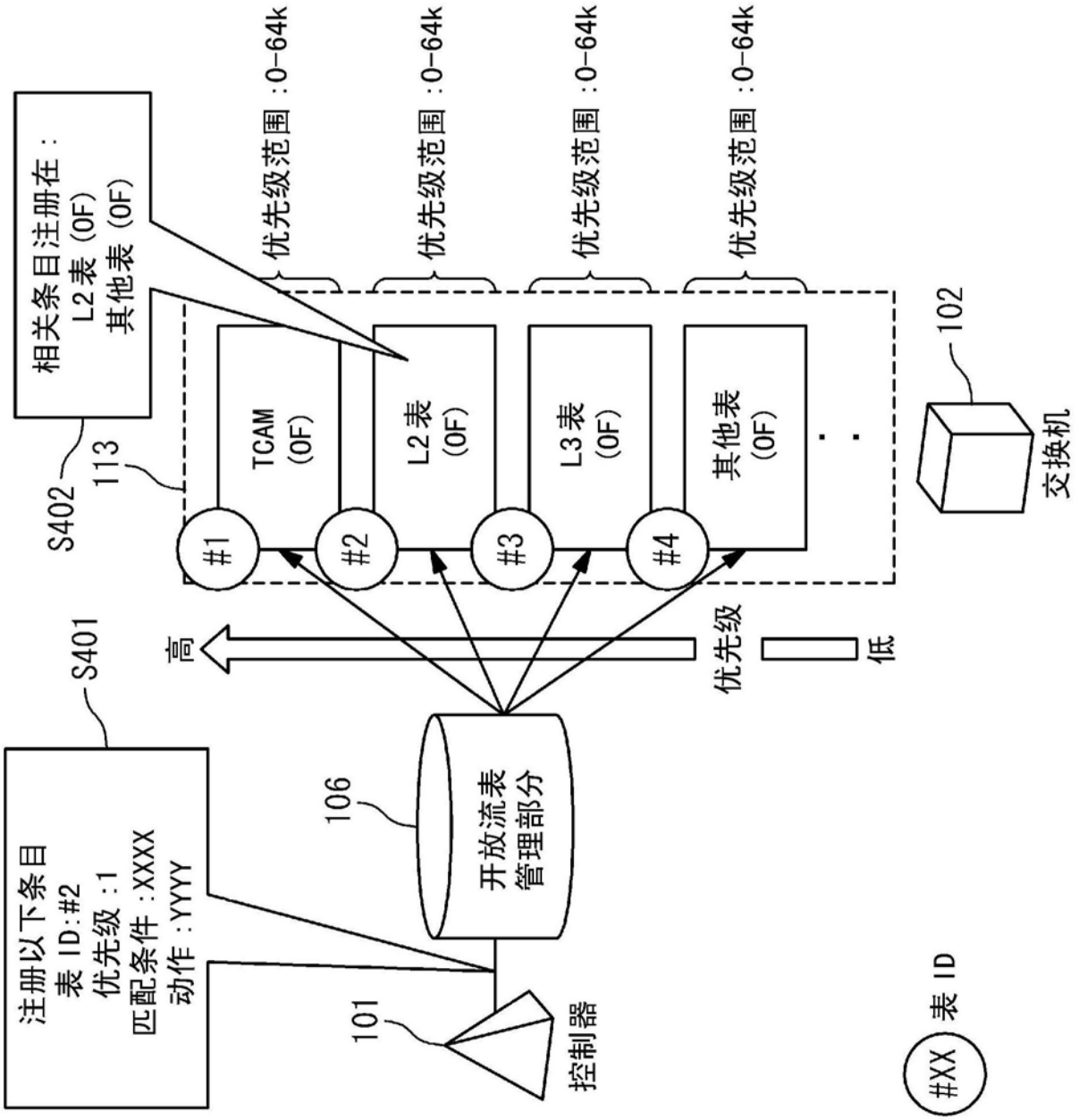

相关条目注册在：
L2表（OF）
其他表（OF）
S402
113
优先级范围：0-64k
优先级范围：0-64k
优先级范围：0-64k
优先级范围：0-64k
TCAM
（OF）
L2表
（OF）
L3表
（OF）
其他表
（OF）
#1
#2
#3
#4
102
交换机
高
优先级
低
S401
注册以下条目
表ID:#2
优先级：1
匹配条件：XXXX
动作：YYYY
106
开放流表
管理部分
101
控制器
#XX 表ID

图10

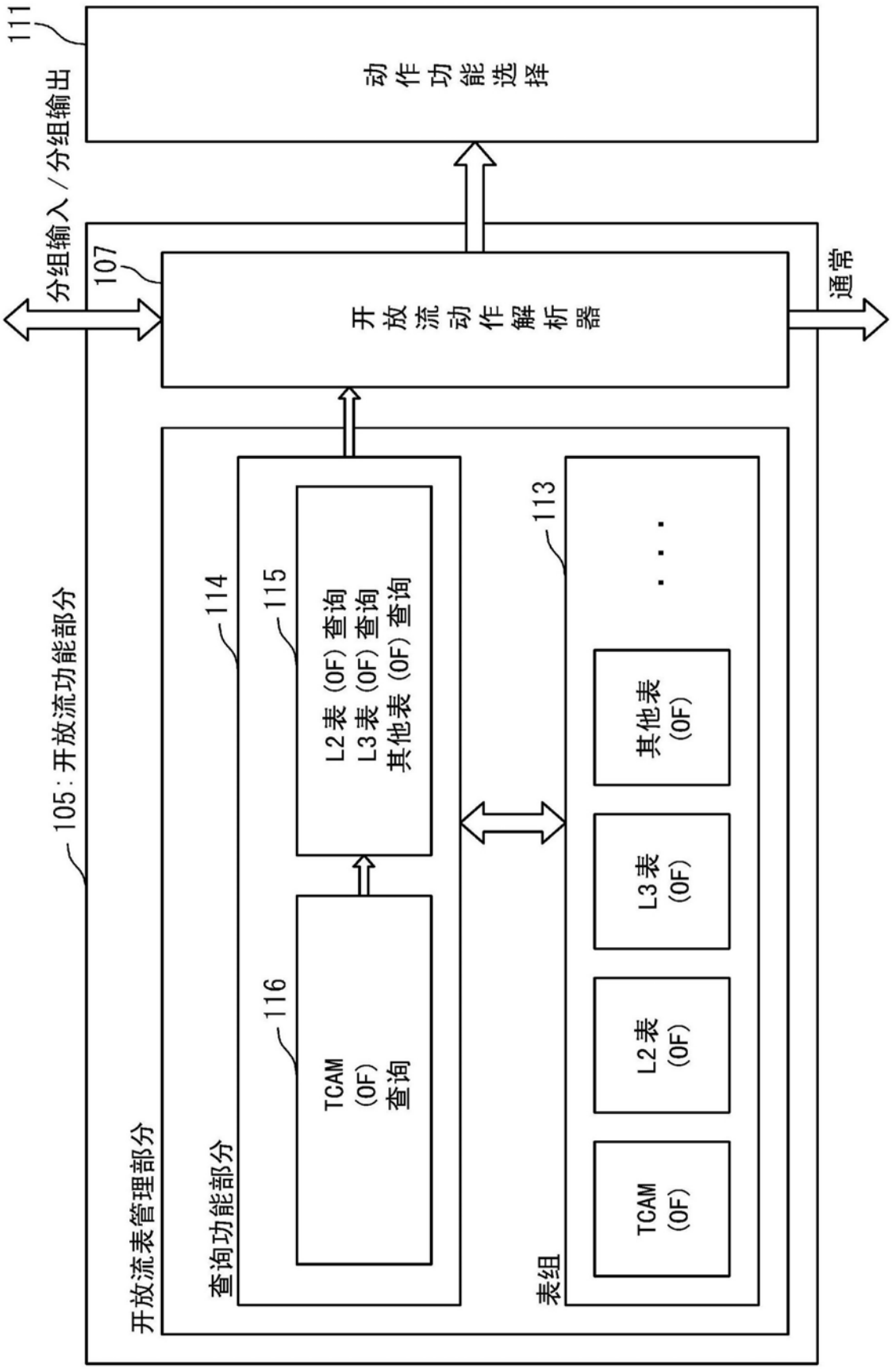

111
动作功能选择
分组输入／分组输出
107
开放流动作解析器
通常
105：开放流功能部分
开放流表管理部分
查询功能部分
114
115
L2表（OF）查询
L3表（OF）查询
其他表（OF）查询
116
TCAM
（OF）
查询
113
表组
TCAM
（OF）
L2表
（OF）
L3表
（OF）
其他表
（OF）
· · ·

图11

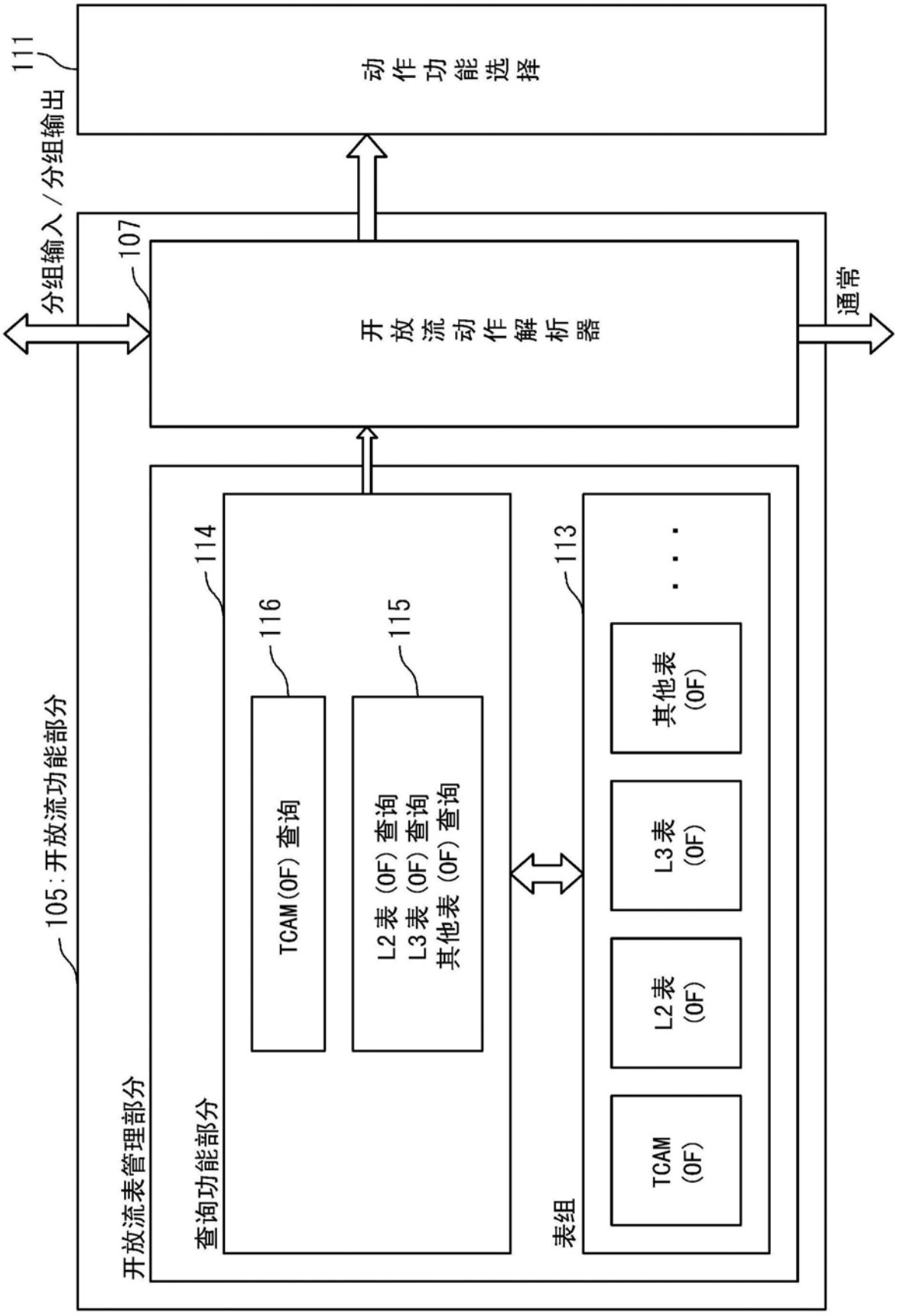

111
动作功能选择
分组输入/分组输出
107
开放流动作解析器
通常
105:开放流功能部分
开放流表管理部分
查询功能部分
114
116
TCAM(OF)查询
115
L2表(OF)查询
L3表(OF)查询
其他表(OF)查询
113
表组
TCAM(OF)
L2表(OF)
L3表(OF)
其他表(OF)
. . .

图12

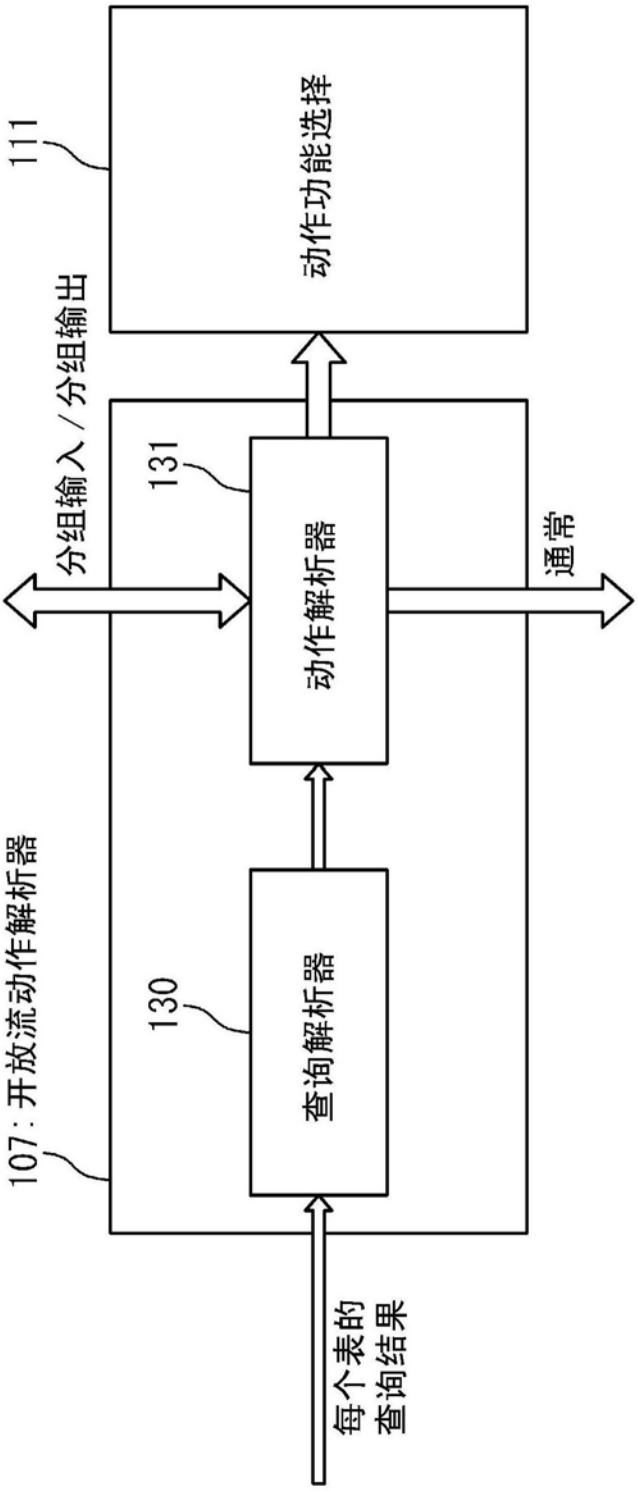

111
动作功能选择
分组输入／分组输出
131
动作解析器
通常
107：开放流动作解析器
130
查询解析器
每个表的
查询结果

图13