



(12) 发明专利申请

(10) 申请公布号 CN 104221331 A

(43) 申请公布日 2014. 12. 17

(21) 申请号 201380017811. 6

(74) 专利代理机构 中国国际贸易促进委员会专利商标事务所 11038

(22) 申请日 2013. 03. 19

代理人 冯玉清

(30) 优先权数据

13/438, 794 2012. 04. 03 US

(51) Int. Cl.

H04L 12/28(2006. 01)

(85) PCT国际申请进入国家阶段日

2014. 09. 29

(86) PCT国际申请的申请数据

PCT/IB2013/052168 2013. 03. 19

(87) PCT国际申请的公布数据

W02013/150404 EN 2013. 10. 10

(71) 申请人 国际商业机器公司

地址 美国纽约

(72) 发明人 K·G·坎博 V·潘德伊 吕达人

S·安南萨拉姆 N·马克杰

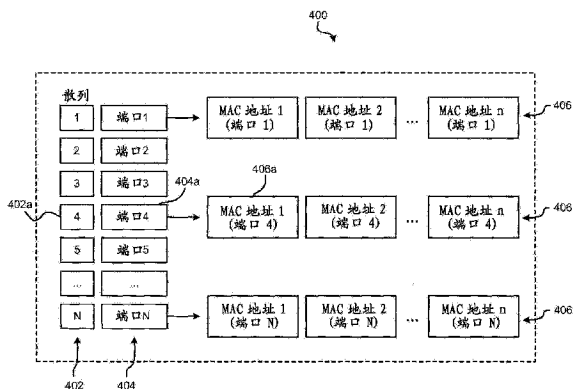
权利要求书3页 说明书9页 附图5页

(54) 发明名称

用于以太网交换机的没有查找表的第2层分组交换

(57) 摘要

在一个实施例中,一种系统包括至少一个处理器,其包括:被配置成接收针对为一个端口上的设备指派介质接入控制(MAC)地址的请求的逻辑;被配置成至少部分地基于所述端口确定将为所述设备指派的MAC地址的逻辑;以及被配置成发送带有MAC地址的针对所述请求的响应的逻辑。在另一个实施例中,一种用于指派MAC地址的计算机程序产品包括其中具体实现有计算机可读程序代码的计算机可读存储介质,所述计算机可读程序代码包括:被配置成在不使用查找表的情况下确定将为设备指派的MAC地址的计算机可读程序代码;以及被配置成向设备发送MAC地址的计算机可读程序代码。根据更多实施例还给出了其他系统、方法和计算机程序产品。



1. 一种包括至少一个处理器的系统,其中,所述至少一个处理器包括:
被配置成接收针对为端口上的设备指派介质接入控制 MAC 地址的请求的逻辑;
被配置成至少部分地基于所述端口确定将为所述设备指派的 MAC 地址的逻辑;以及
被配置成发送带有 MAC 地址的针对所述请求的响应的逻辑。
2. 如权利要求 1 所述的系统,还包括被配置成创建 MAC 地址分配表的逻辑,
其中,所述 MAC 地址分配表包括多个散列数值,每一个散列数值与一个端口和多个 MAC 地址相关联,其中对于所述多个 MAC 地址当中的任一个应用散列算法会得到相关联的散列数值。
3. 如权利要求 2 所述的系统,其中,被配置成为设备指派 MAC 地址的逻辑包括被配置成从 MAC 地址分配表中选择 MAC 地址的逻辑,从而使得对所选 MAC 地址应用散列算法会得到与耦合到所述设备的端口相关联的散列数值。
4. 如权利要求 3 所述的系统,其中,被配置成从 MAC 地址分配表中选择 MAC 地址的逻辑包括被配置成从多个可用 MAC 地址当中顺序地或随机地选择下一个可用 MAC 地址的逻辑,其中可用 MAC 地址当前未被指派给设备。
5. 如权利要求 4 所述的系统,其还包括:
在从 MAC 地址分配表中选择 MAC 地址时,通过将所述 MAC 地址从 MAC 地址分配表中去除并且将所述 MAC 地址添加到已使用 MAC 地址表中,或者通过在 MAC 地址分配表中翻转与所述 MAC 地址相关联的指示比特而被配置成将所述 MAC 地址标记为不可用的逻辑;以及
在设备释放为其指派的 MAC 地址时,通过将所述 MAC 地址添加到 MAC 地址分配表中并且从已使用 MAC 地址表中去除所述 MAC 地址,或者通过在 MAC 地址分配表中翻转与所述 MAC 地址相关联的指示比特而被配置成将所述 MAC 地址标记为可用的逻辑。
6. 如权利要求 3 所述的系统,其中,所述散列算法是: $H(DM) = (DM[0]*1XOR DM[1]*2XOR DM[2]*3XOR DM[3]*4XOR DM[4]*5XOR DM[5]*6) MOD P$,其中 DM 是 MAC 地址, H(DM) 是 MAC 地址 DM 的所有六个字节的散列数值,并且 P 是系统的端口总数。
7. 如权利要求 2 所述的系统,其中,在所述至少一个处理器的启动处理期间创建 MAC 地址分配表,并且其中所述散列算法包括: $H(DM) = (DM[0]*1XOR DM[1]*2... XOR DM[N-1]*N) MOD P$,其中 DM 是 MAC 地址, H(DM) 是 MAC 地址 DM 的散列数值, DM[N-1] 是具有 N 个字节的 MAC 地址 DM 的字节,并且 P 是端口总数。
8. 如权利要求 1 所述的系统,其中,所述请求包括具有源地址的 MAC 地址采集协议 MAAP 分组,并且其中所述响应包括经过改动的所接收到的 MAAP 分组,从而使得经过改动的 MAAP 分组的目的地地址被改变到所接收到的 MAAP 分组的源地址。
9. 一种用于指派介质接入控制 MAC 地址的计算机程序产品,所述计算机程序产品包括其中具体实现有计算机可读程序代码的计算机可读存储介质,所述计算机可读程序代码包括:
被配置成在不使用查找表的情况下确定将为设备指派的 MAC 地址的计算机可读程序代码;以及
被配置成向设备发送 MAC 地址的计算机可读程序代码。
10. 如权利要求 9 所述的计算机程序产品,还包括:
被配置成创建 MAC 地址分配表的计算机可读程序代码,其中所述 MAC 地址分配表包括

多个散列数值,每一个散列数值与对应于多个 MAC 地址的端口相关联,对所述多个 MAC 地址当中的任一个应用散列算法会得到相关联的散列数值;以及

被配置成从 MAC 地址分配表中的多个 MAC 地址当中选择 MAC 地址的计算机可读程序代码,从而使得对所选 MAC 地址应用散列算法会得到与耦合到所述设备的端口相关联的散列数值。

11. 如权利要求 10 所述的计算机程序产品,还包括,在从 MAC 地址分配表中选择 MAC 地址时,被配置成将所述 MAC 地址标记为不可用的计算机可读程序代码,或者被配置成将所述 MAC 地址从 MAC 地址分配表中去除并且将所述 MAC 地址添加到已使用 MAC 地址表中的计算机可读程序代码,其中不可用 MAC 地址当前被指派给设备。

12. 如权利要求 11 所述的计算机程序产品,还包括在设备释放为其指派的 MAC 地址时:

被配置成将所述 MAC 地址添加到 MAC 地址分配表中并且从已使用 MAC 地址表中去除所述 MAC 地址的计算机可读程序代码;或者

被配置成将所述 MAC 地址标记为可用的计算机可读程序代码。

13. 如权利要求 12 所述的计算机程序产品,其中,被配置成将 MAC 地址标记为可用的计算机可读程序代码包括被配置成在 MAC 地址分配表中翻转与所述 MAC 地址相关联的指示比特的逻辑。

14. 如权利要求 10 所述的计算机程序产品,其中,所述散列算法是: $H(DM) = (DM[0]*1XOR DM[1]*2XOR DM[2]*3XOR DM[3]*4XOR DM[4]*5XOR DM[5]*6)MOD P$,其中 DM 是 MAC 地址, H(DM) 是 MAC 地址 DM 的所有六个字节的散列数值,并且 P 是端口总数。

15. 如权利要求 10 所述的计算机程序产品,其中,在启动处理期间创建 MAC 地址分配表,并且所述散列算法包括: $H(DM) = (DM[0]*1XOR DM[1]*2 \dots XOR DM[N-1]*N)MOD P$,其中 DM 是 MAC 地址, H(DM) 是 MAC 地址 DM 的散列数值, DM[N-1] 是具有 N 个字节的 MAC 地址 DM 的字节,并且 P 是端口总数。

16. 一种用于指派介质接入控制 MAC 地址的方法,所述方法包括:

在交换设备的端口处接收来自设备的针对 MAC 地址的请求;

至少部分地基于所述端口确定将为所述设备指派的 MAC 地址;以及

利用所述 MAC 地址对所述请求做出响应。

17. 如权利要求 16 所述的方法,还包括创建 MAC 地址分配表,其中所述 MAC 地址分配表包括多个散列数值,每一个散列数值与对应于多个 MAC 地址的端口相关联,对所述多个 MAC 地址当中的任一个应用散列算法会得到相关联的散列数值。

18. 如权利要求 17 所述的方法,其中, MAC 地址分配服务器 MAAS 使用从输入文件提供到 MAAS 的预先定义的 MAC 地址集合来建立 MAC 地址分配表,其中 MAAS 响应于接收到 MAC 地址分配协议 MAAP 分组从 MAC 地址分配表中确定将为设备指派的 MAC 地址。

19. 如权利要求 17 所述的方法,其中,确定将为设备指派的 MAC 地址包括从 MAC 地址分配表中选择所述 MAC 地址,从而使得对所选 MAC 地址应用散列算法会得到与耦合到所述设备的交换设备的端口相关联的散列数值。

20. 如权利要求 18 所述的方法,其中,从 MAC 地址分配表中选择 MAC 地址包括从多个可用 MAC 地址当中顺序地或随机地选择下一个可用 MAC 地址,其中可用 MAC 地址是当前未被

指派给设备的 MAC 地址。

21. 如权利要求 18 所述的方法,其还包括:

在从 MAC 地址分配表中选择 MAC 地址时,将所述 MAC 地址标记为不可用或者从 MAC 地址分配表中去除所述 MAC 地址并且将所述 MAC 地址添加到已使用 MAC 地址表中;以及

在设备释放为其指派的 MAC 地址时,将所述 MAC 地址添加到 MAC 地址分配表中并且从已使用 MAC 地址表中去除所述 MAC 地址或者将所述 MAC 地址标记为可用。

22. 如权利要求 18 所述的方法,其中,所述散列算法是: $H(DM) = (DM[0]*1XOR DM[1]*2XOR DM[2]*3XOR DM[3]*4XOR DM[4]*5XOR DM[5]*6)MOD P$,其中 DM 是 MAC 地址, H(DM) 是 MAC 地址 DM 的所有六个字节的散列数值,并且 P 是交换设备的端口总数。

23. 如权利要求 17 所述的方法,其中,在所述交换设备的启动处理期间创建 MAC 地址分配表,并且所述散列算法包括: $H(DM) = (DM[0]*1XOR DM[1]*2... XOR DM[N-1]*N)MOD P$,其中 DM 是 MAC 地址, H(DM) 是 MAC 地址 DM 的散列数值, DM[N-1] 是具有 N 个字节的 MAC 地址 DM 的字节,并且 P 是端口总数。

24. 一种用于取回介质接入控制 MAC 地址的方法,所述方法包括:

向 MAC 分配服务器 MAAS 发送针对 MAC 地址的请求;

等待预定时间量以接收针对所述请求的响应,其中所述响应包括 MAC 地址;以及在所述预定时间量内接收到针对所述请求的响应时使用所述 MAC 地址。

25. 如权利要求 24 所述的方法,其还包括在所述预定时间量内未接收到响应时使用预先编程的 MAC 地址。

用于以太网交换机的没有查找表的第 2 层分组交换

背景技术

[0001] 本发明涉及数据中心基础设施,更具体来说,本发明涉及能够在不使用查找表的情况下进行第 2 层分组交换的以太网交换机。

[0002] 在分组交换网络交换机中,通常对目的地介质接入控制 (MAC) 地址或目的地互联网协议 (IP) 地址实施查找,以便发现用于转发分组的交换设备的实际目的地端口。大多数系统利用某种交换处理器来进行这一确定,比如专用集成电路 (ASIC)。所述交换处理器保持用于实施查找解析的查找表,其在典型的网络中可能相当长,并且可能具有数以千计的条目。所述查找表被存储在芯片上或者被存储在芯片外部。在任一种情况下,对于查找表的使用都会对于芯片增加显著的逻辑复杂度、空间和存储器需求。这一附加的复杂度最终会增加更多等待时间,并且使得芯片比起所期望的情况更加昂贵。因此,有益的是将能够在不使用查找表的情况下发现实际目的地端口。

发明内容

[0003] 在一个实施例中,一种系统包括至少一个处理器,其包括:被配置成接收针对为一个端口上的设备指派介质接入控制 (MAC) 地址的请求的逻辑;被配置成至少部分地基于所述端口确定将为所述设备指派的 MAC 地址的逻辑;

[0004] 以及被配置成发送带有 MAC 地址的针对所述请求的响应的逻辑。

[0005] 在另一个实施例中,一种用于指派 MAC 地址的计算机程序产品包括随之具体实现计算机可读程序代码的计算机可读存储介质,所述计算机可读程序代码包括:被配置成在不使用查找表的情况下确定将为设备指派的 MAC 地址的计算机可读程序代码;以及被配置成向设备发送 MAC 地址的计算机可读程序代码。

[0006] 根据另一个实施例,一种用于指派 MAC 地址的方法包括:在交换设备的端口处接收来自设备的针对 MAC 地址的请求;至少部分地基于所述端口确定将为所述设备指派的 MAC 地址;以及利用所述 MAC 地址对所述请求做出响应。

[0007] 在另一个实施例中,一种用于取回 MAC 地址的方法包括:向 MAC 分配服务器 (MAAS) 发送针对 MAC 地址的请求;等待预定时间量以接收针对所述请求的响应,所述响应包括 MAC 地址;以及在所述预定时间量内接收到针对所述请求的响应时使用所述 MAC 地址。

[0008] 通过后面结合附图做出的以举例的方式说明本发明的原理的详细描述,本发明的其他方面和实施例将变得显而易见。

附图说明

[0009] 图 1 示出了根据一个实施例的网络架构。

[0010] 图 2 示出了根据一个实施例的可以与图 1 的服务器和 / 或客户端相关联的代表性硬件环境。

[0011] 图 3 是根据一个实施例的交换设备的简化图示。

[0012] 图 4 示出了根据一个实施例的介质接入控制 (MAC) 地址分配表。

[0013] 图 5 是根据一个实施例的方法的流程图。

[0014] 图 6 是根据另一个实施例的方法的流程图。

具体实施方式

[0015] 后面的描述是为了说明本发明的一般原理,而不意图限制这里所要求保护的发明性概念。此外,这里所描述的特定特征可以与多种可能的组合和置换当中的每一种组合和置换中的其他所描述的特征组合使用。

[0016] 除非在这里明确地另行定义,否则应当为所有术语给出其所可能的最宽泛的解释,其中包括说明书所暗指的含义以及本领域技术人员所理解和 / 或如词典、协定中所定义的含义。

[0017] 还必须提到的是,除非另行表明,否则用在说明书和所附权利要求书中的单数形式“一个”、“一项”和“所述”也包括复数的所指对象。

[0018] 在一种方法中,可以对于耦合到交换系统的遵循某一系统的设备确定介质接入控制 (MAC) 地址,其中每一个 MAC 地址在被应用散列算法时对应于特定散列数值,从而使得每一个散列数值对应于交换设备的一个端口,从而简化并且改进为设备指派 MAC 地址的方式。

[0019] 在一个一般性实施例中,一种系统包括至少一个处理器,其包括:被配置成接收针对为一个端口上的设备指派介质接入控制 (MAC) 地址的请求的逻辑;被配置成至少部分地基于所述端口确定将为所述设备指派的 MAC 地址的逻辑;以及被配置成发送带有 MAC 地址的针对所述请求的响应的逻辑。

[0020] 在另一个一般性实施例中,一种用于指派 MAC 地址的计算机程序产品包括随之具体实现计算机可读程序代码的计算机可读存储介质,所述计算机可读程序代码包括:被配置成在不使用查找表的情况下确定将为设备指派的 MAC 地址的计算机可读程序代码;以及被配置成向设备发送 MAC 地址的计算机可读程序代码。

[0021] 根据另一个一般性实施例,一种用于指派 MAC 地址的方法包括:在交换设备的端口处接收来自设备的针对 MAC 地址的请求;至少部分地基于所述端口确定将为所述设备指派的 MAC 地址;以及利用所述 MAC 地址对所述请求做出响应。

[0022] 在另一个一般性实施例中,一种用于取回 MAC 地址的方法包括:向 MAC 分配服务器 (MAAS) 发送针对 MAC 地址的请求;等待预定时间量以接收针对所述请求的响应,所述响应包括 MAC 地址;以及在所述预定时间量内接收到针对所述请求的响应时使用所述 MAC 地址。

[0023] 所属技术领域的技术人员知道,本发明的各个方面可以实现为系统、方法或计算机程序产品。因此,本发明的各个方面可以具体实现为以下形式,即:完全的硬件实施方式、完全的软件实施方式(包括固件、驻留软件、微代码等),或硬件和软件方面结合的实施方式,这里可以统称为“电路”、“模块”或“系统”。此外,在一些实施例中,本发明的各个方面还可以实现为在一个或多个计算机可读介质中的计算机程序产品的形式,该计算机可读介质中包含计算机可读的程序代码。

[0024] 可以利用一个或多个计算机可读介质的任意组合。计算机可读介质可以是计算机可读信号介质或者非瞬时性计算机可读存储介质。非瞬时性计算机可读存储介质例如可以是——但不限于——电子、磁、光、电磁、红外线、或半导体的系统、装置或设备,或者任意以

上的组合。非瞬时性计算机可读存储介质的更具体的例子（非穷举的列表）包括：便携式计算机盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、便携式紧凑盘只读存储器（CD-ROM）、蓝光盘只读存储器（BD-ROM）、光存储设备、磁存储设备、或者上述的任意合适的组合。在本文件中，非瞬时性计算机可读存储介质可以是任何能够包含或存储程序或应用的有形介质，该程序或应用供指令执行系统、装置或者设备使用或者与其结合使用。

[0025] 计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括——但不限于——电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者设备使用或者与其结合使用的程序，比如具有一条或多条连线的电连接、光纤等等。

[0026] 计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括——但不限于——无线、有线、光缆、RF 等等，或者上述的任意合适的组合。

[0027] 可以以一种或多种程序设计语言的任意组合来编写用于执行本发明操作的计算机程序代码，所述程序设计语言包括面向对象的程序设计语言——诸如 Java、Smalltalk、C++ 等，还包括常规的过程式程序设计语言——诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网（LAN）存储区域网（SAN）和 / 或广域网（WAN）——连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

[0028] 下面将参照根据本发明实施例的方法、装置（系统）和计算机程序产品的流程图和 / 或框图描述本发明。应当理解，流程图和 / 或框图的每个方框以及流程图和 / 或框图中各方框的组合，都可以由计算机程序指令实现。这些计算机程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器，从而生产出一种机器，使得这些计算机程序指令在通过计算机或其它可编程数据处理装置的处理器执行时，产生了实现流程图和 / 或框图中的一个或多个方框中规定的功能 / 动作的装置。

[0029] 也可以把这些计算机程序指令存储在计算机可读介质中，这些指令使得计算机、其它可编程数据处理装置、或其他设备以特定方式工作，从而，存储在计算机可读介质中的指令就产生出包括实现流程图和 / 或框图中的一个或多个方框中规定的功能 / 动作的指令的制品（article of manufacture）。

[0030] 计算机程序指令还可以被加载到计算机、其他可编程数据处理装置或者其他设备上，从而使得在所述计算机、其他可编程装置或其他设备上实施一系列操作步骤以产生计算机实施的处理，从而使得在计算机或其他可编程装置上执行的指令提供用于实施在流程图和 / 或方框图的一个或多个方框中规定的功能 / 动作的处理。

[0031] 图 1 示出了根据一个实施例的网络架构 100。如图 1 中所示，提供多个远程网络 102，其中包括第一远程网络 104 和第二远程网络 106。网关 101 可以耦合在远程网络 102 与邻近网络 108 之间。在这一网络架构 100 的情境中，网络 104、106 可以分别采取任何形

式,其中包括但不限于 LAN、例如因特网之类的 WAN、公共交换电话网 (PSTN)、内部电话网等等。

[0032] 在使用中,网关 101 充当从远程网络 102 到邻近网络 108 的进入点。因此,网关 101 可以充当能够引导到达该网关 101 处的给定数据分组的路由器,以及对于给定分组布置进入和离开该网关 101 的实际路径的交换机。

[0033] 此外还包括耦合到邻近网络 108 的至少一个数据服务器 114,其可以从远程网络 102 经由网关 101 访问。应当提到的是,(多个)数据服务器 114 可以包括任何类型的计算设备/群件。每一个数据服务器 114 耦合有多个用户设备 116。这样的用户设备 116 可以包括台式计算机、膝上型计算机、手持式计算机、打印机以及/或者其他类型的包含逻辑的设备。应当提到的是,在某些实施例中,用户设备 111 还可以直接耦合到任何网络。

[0034] 一个外设 120 或一系列外设 120 可以耦合到一个或多个网络 104、106、108,所述外设比如是传真机、打印机、扫描仪、硬盘驱动器、联网和/或本地存储单元或系统等等。应当提到的是,数据库和/或附加的组件可以与耦合到网络 104、106、108 的任何类型的网络设备一同使用或者被集成到其中。在本发明的描述的情境中,网络元件可以指代网络的任何组件。

[0035] 根据某些方法,这里所描述的方法和系统可以利用虚拟系统和/或仿真一个或多个其他系统的系统来实施以及/或者被实施在其上,比如仿真 IBM z/OS 环境的 UNIX 系统、虚拟主机 MICROSOFT WINDOWS 环境的 UNIX 系统、仿真 IBM z/OS 环境的 MICROSOFT WINDOWS 系统等等。在一些实施例中,通过使用 VMWARE 软件可以增强这一虚拟化和/或仿真。

[0036] 在更多方法中,一个或多个网络 104、106、108 可以代表被共同称为“云”的系统群集。在云计算中,通过按需关系为云中的任何系统提供例如处理能力、外设、软件、数据、服务器等共享资源,从而允许在许多计算系统之间访问和分配服务。云计算通常涉及在云中操作的系统之间的因特网连接,但是也可以采用本领域内已知的其他连接系统的技术。

[0037] 图 2 示出了根据一个实施例的与图 1 的用户设备 116 和/或服务器 114 相关联的代表性硬件环境。图 2 示出了工作站的典型硬件配置,其具有中央处理单元 (CPU) 210 (比如微处理器) 以及经由一条或多条总线 212 互连的若干其他单元,所述总线根据几个实施例可以是不同类型,比如局部总线、并行总线、串行总线等等。

[0038] 图 2 中所示出的工作站包括随机存取存储器 (RAM) 214,只读存储器 (ROM) 216,用于把例如盘存储单元 220 之类的外围设备连接到一条或多条总线 212 的 I/O 适配器 218,用于把键盘 224、鼠标 226、扬声器 228、麦克风 232 和/或例如触摸屏、数字摄影机(未示出)等其他用户接口设备连接到一条或多条总线 212 的用户接口适配器 222,用于把所述工作站连接到通信网络 235 (例如数据处理网络) 的通信适配器 234,以及用于把一条或多条总线 212 连接到显示设备 238 的显示适配器 236。

[0039] 所述工作站上可以驻留有操作系统,比如 MICROSOFT WINDOWS 操作系统 (OS)、MAC OS、UNIX OS 等等。应当认识到,一个优选实施例还可以被实施在所提到的之外的其他平台和操作系统。一个优选实施例可以利用 JAVA、XML、C 和/或 C++ 语言或其他编程语言连同面向对象的编程方法来编写。可以使用越来越多地被用来开发复杂应用的面向对象的编程 (OOP)。

[0040] 当前,IEEE 向例如公司、教育机构等组织指派 MAC 地址,其随后按照需要在其(多

个)网络上自由指派 MAC 地址。但是这一处理效率低下,并且存在大量等待时间和问题。举例来说,在以太网交换中,取决于存储对于特定交换设备所获知的所有 MAC 地址的 MAC 地址表的尺寸,在对交换机进行扩展以用在不断扩张的网络中时,在针对 MAC 地址解析的处理能力和保持较低查找等待时间方面可能存在很大问题。

[0041] 为了克服与当前使用的 MAC 编址协议相关联的问题和等待时间,取代指派固定 MAC 地址,可以在耦合到要求 MAC 地址的设备的交换设备引导 (boot-up) 时指派 MAC 地址。这些 MAC 地址可以基于对算法分配的符合来确定。

[0042] 根据一个实施例,将把所接收到的分组交换到该处的目的地端口可以是目的地 MAC 地址的函数,其可以被写作 $Dest_Port = F(Dest_MAC_Address, is_mirror, is_flow)$, 其中 F 表示函数, Dest_Port 是目的地端口, Dest_MAC_Address 是由分组提供的目的地 MAC 地址, is_mirror 表示该分组是否被允许镜像到多个端口,并且 is_flow 表示特殊操作和流控制。is_mirror 和 is_flow 对于分组交换不是必须的,但是在希望时可以使用。这样,与传统方法相比,MAC 地址分配被更加高效地实施并且所需的开销处理更少。

[0043] 现在参照图 3,其中示出了根据一个实施例的系统 300,比如交换设备。如图所示,系统 300 包括服务器,比如 MAAS,其被适配成对于电耦合到系统 300 的设备 304 指派和管理 MAC 地址。所述服务器可以是一个组件并且 / 或者由至少一个处理器 306 管理,比如用于执行逻辑的交换处理器。在一些实施例中,系统 300 还可以包括用于执行逻辑的本地处理器 302,其电耦合到所述至少一个处理器 306。对于本地处理器 302 和 / 或至少一个处理器 306 可以使用本领域内已知的任何类型的处理器,比如中央处理单元 (CPU)、现场可编程门阵列 (FPGA)、集成电路 (IC)、专用集成电路 (ASIC) 等等。在另一个实施例中,交换 ASIC 和本地 CPU 都可以被包括在系统 300 中,其中交换 ASIC 管理交换操作,本地 CPU 管理 MAC 地址分配操作。

[0044] 每一个变为在线的新的主机、服务器或设备采集 MAC 地址以便与其他设备通信。这从传统方法来说是唯一的,其中设备的网络接口卡 (NIC) 预先编程有 MAC 地址。在传统上,设备在引导时使用该 MAC 地址进行通信。但是根据这里所描述的实施例,一种利用 MAC 地址采集算法或 MAC 地址采集协议 (MAAP) 的 MAC 地址采集方法不再需要预先编程的 MAC 地址,而是可以在某些方法中作为针对预先编程的 MAC 地址的补充来使用。

[0045] 根据一个实施例,所述至少一个处理器 306 包括或者可以访问被配置成接收针对为端口 308 上的设备 304 指派 MAC 地址的请求的逻辑。所述请求可以由设备 304 发送,或者可以源自网络中或系统 300 中的其他地方。在该处接收到请求的端口 308 被存储或者以其他方式被记住,从而使其可以被用来在将来发送去往设备 304 的所有外出通信量。所述至少一个处理器 306 还包括或者可以访问被配置成至少部分地基于所述端口确定将为设备指派的 MAC 地址的逻辑,以及被配置成发送带有 MAC 地址的针对所述请求的响应的逻辑。

[0046] 在一个实施例中,所述至少一个处理器 306 还可以包括或者可以访问被配置成创建 MAC 地址分配表的逻辑。现在参照图 4,MAC 地址分配表 400 可以包括多个散列数值 402,每一个散列数值 402 与一个端口 404 和多个 MAC 地址 406 相关联。在一种方法中,现在参照图 3-4,对所述多个 MAC 地址 406 当中的任一个应用散列算法会得到相关联的散列数值 402,从而存在一种为设备 304 指派 MAC 地址 406 的系统性方法,其中设备 304 请求对应于该设备 304 在其上耦合到系统 300 的端口 404 的 MAC 地址 406。

[0047] 在一种方法中,被配置成确定将为设备 304a 指派的 MAC 地址 406a 的逻辑可以包括被配置成从 MAC 地址分配表 400 中选择 MAC 地址 406a 的逻辑,从而使得对所选 MAC 地址 406a 应用散列算法会得到与耦合到设备 304a 的端口 404a 相关联的散列数值 402a。

[0048] 在另一种方法中,被配置成从 MAC 地址分配表 400 中选择 MAC 地址 406a 的逻辑包括被配置成从多个可用 MAC 地址 406 当中顺序地或随机地选择下一个可用 MAC 地址 406 的逻辑,其中可用 MAC 地址当前未被分配给设备 304。当然,与此相反,不可用 MAC 地址当前被指派给设备 304。

[0049] 在另一种方法中,所述逻辑还可以包括在从 MAC 地址分配表 400 中选择 MAC 地址 406a 时被配置成将 MAC 地址 406a 标记为不可用的逻辑,这是通过将 MAC 地址 406a 从 MAC 地址分配表 400 中去除并且将所述 MAC 地址添加到已使用 MAC 地址表(未示出,但是其将类似于 MAC 地址分配表 400,不同之处在于其将仅包括已指派 MAC 地址)中。在更多方法中,所述逻辑还可以包括在设备 304a 释放为之指派的 MAC 地址 406a 时被配置成将 MAC 地址 406a 标记为可用的逻辑,这是通过将 MAC 地址 406a 添加到 MAC 地址分配表 400 中并且从已使用 MAC 地址表 400 中去除 MAC 地址 406a。

[0050] 在另一个实施例中,所述逻辑可以被配置成在 MAC 地址分配表 400 中翻转与 MAC 地址 406a 相关联的指示比特,从而表明 MAC 地址 406a 可用或不可用。

[0051] 根据各个实施例,所述散列算法可以包括本领域内已知的任何完美或不完美散列算法。在一种方法中,所述散列算法可以包括: $H(DM) = (DM[0]*1XOR DM[1]*2XOR DM[2]*3XOR DM[3]*4XOR DM[4]*5XOR DM[5]*6)MOD P$,其中 DM 是 MAC 地址,H(DM) 是 MAC 地址 DM 的所有六个字节的散列数值,并且 P 是系统 300 的物理端口 308 的总数。

[0052] 在另一个实施例中,所述散列算法可以包括: $H(DM) = (DM[0]*1XOR DM[1]*2... XOR DM[N-1]*N)MOD P$,其中 DM 是 MAC 地址,H(DM) 是 MAC 地址 DM 的散列数值,DM[N-1] 是具有 N 个字节的 MAC 地址 DM 的一个字节,并且 P 是物理端口总数。

[0053] 根据另一个实施例,可以在所述至少一个处理器 306 的启动(start-up)处理期间创建 MAC 地址分配表 400。这一启动处理可以是系统 300 引导,或者是发生一次以便令各个组件达到运转速度的任何其他处理。

[0054] 在另一个实施例中,所述请求可以包括具有源地址的 MAC 地址采集协议(MAAP)分组。在这种情况下,所述响应可以包括经过改动的所接收到的 MAAP 分组,从而使得经过改动的 MAAP 分组的地址被改变到所接收到的 MAAP 分组的源地址。这样,通过源和目的地地址之间的简单切换,相同的 MAAP 分组可以被用于请求和响应。

[0055] 现在参照图 5,其中示出了根据一个实施例的用于指派 MAC 地址的方法 500 的流程图。在各个实施例中,特别可以根据本发明在图 1-4 所描绘的任何环境中实施方法 500。当然,本领域技术人员通过阅读本发明的描述将会理解的是,在方法 500 中可以包括比在图 5 中具体描述的那些操作更多或更少的操作。

[0056] 方法 500 的每一个步骤可以由操作环境的任何适当组件实施。举例来说,在一个实施例中,方法 500 可以部分地或完全地由交换设备和/或交换设备的 MAAS 实施。

[0057] 如图 5 中所示,方法 500 可以发起于操作 502,其中在交换设备的一个端口处接收来自设备的针对 MAC 地址的请求。正如这里所描述的那样,在该处接收到请求的端口被存储以用于确定对应于所述设备的 MAC 地址。

[0058] 在操作 504 中,至少部分地基于所述端口确定将为设备指派的 MAC 地址。

[0059] 在一种方法中,可以创建 MAC 地址分配表。MAC 地址分配表可以包括多个散列数值,每一个散列数值与对应于多个 MAC 地址的一个端口相关联。此外,对所述多个 MAC 地址当中的任一个应用散列算法会得到相关联的散列数值,其又与所述 MAC 地址的端口相关联。

[0060] 此外,在一个实施例中,MAAS 可以使用从输入文件提供到 MAAS 的预先定义的 MAC 地址集合来建立 MAC 地址分配表。在这种情况下,MAAS 响应于接收到 MAC 地址分配协议(MAAP)分组从 MAC 地址分配表中确定将为设备指派的 MAC 地址。

[0061] 在一个实施例中,所述散列算法可以是完美散列算法或不完美散列算法。可以使用的一种此类完美散列算法是: $H(DM) = (DM[0]*1XOR DM[1]*2XOR DM[2]*3XOR DM[3]*4XOR DM[4]*5XOR DM[5]*6)MOD P$,其中 DM 是 MAC 地址,H(DM) 是 MAC 地址 DM 的所有六个字节的散列数值,并且 P 是交换设备的物理端口的总数。

[0062] 根据另一个实施例,所述散列算法可以包括 $H(DM) = (DM[0]*1XOR DM[1]*2... XOR DM[N-1]*N)MOD P$,其中 DM 是 MAC 地址,H(DM) 是 MAC 地址 DM 的散列数值,DM[N-1] 是具有 N 个字节的 MAC 地址 DM 的一个字节,并且 P 是端口总数。

[0063] 在某些方法中,如果接收到来自设备的分组并且对 MAC 地址应用散列算法没有得到交换设备的适当端口,例如目的地 MAC 地址并非由 MAAP 提供,则可以使用常规查找算法来找到目的地端口。随后可以将所述分组转发到目的地端口。为了识别出并非由 MAAS 利用 MAAP 提供的 MAC 地址,可以比较 MAC 地址的全球唯一标识符(UUID)字段的范围。MAAP 在特定预先编程的范围内分配 MAC 地址。

[0064] 在一种方法中,可以在交换设备的启动处理期间(比如在引导期间)创建 MAC 地址分配表。此外,MAC 地址分配可以是列表、关联、文件或者某种其他手段,其允许将一个或多个 MAC 地址与单一端口和单一散列数值相关联。

[0065] 在另一种方法中,方法 500 还可以包括从 MAC 地址分配表中选择 MAC 地址,从而使对所选 MAC 地址应用散列算法会得到与耦合到所述设备的交换设备的端口相关联的散列数值。举例来说,如果设备连接到交换设备的端口 4,则对为所述设备选择的 MAC 地址应用散列算法会得到与端口 4 相关联的散列数值。所述散列数值可以是 4,或者可以是在指派散列到该特定散列数值的 MAC 地址之前与交换设备的所述端口相关联的任何其他数字。

[0066] 在另一种方法中,从 MAC 地址分配表中选择 MAC 地址可以包括从多个可用 MAC 地址当中顺序地或随机地选择下一个可用 MAC 地址。可用 MAC 地址是当前未被指派给设备的 MAC 地址,不可用 MAC 地址则是当前被指派给设备的 MAC 地址。

[0067] 在另一个实施例中,方法 500 可以包括在从 MAC 地址分配表中选择 MAC 地址时,将所述 MAC 地址标记为不可用,或者将所述 MAC 地址从 MAC 地址分配表中去除并且将所述 MAC 地址添加到已使用 MAC 地址表中。此外,在设备释放为之指派的 MAC 地址时,可以将所述 MAC 地址添加到 MAC 地址分配表中并且从已使用 MAC 地址中去除,或者可以将所述 MAC 地址标记为可用。当然,在另一些实施例中可以使用全部两个单独的分配表和标记。

[0068] 在另一种方法中,所述标记可以包括翻转 MAC 地址分配表中的与 MAC 地址相关联的指示比特,或者本领域内已知的对表中的条目进行标记的任何方法。

[0069] 在操作 506 中,利用 MAC 地址对所述请求做出响应。可以使用包括 MAC 地址的任

何响应,比如以太网分组。在一种此类情况中,所述请求可以包括具有源地址的 MAAP 分组。在这种情况下,所述响应可以包括经过改动的所接收到的 MAAP 分组,从而使得经过改动的 MAAP 分组的目的地地址被改变到所接收到的 MAAP 分组的源地址。这样,通过源和目的地地址之间的简单切换,相同的 MAAP 分组可以被用于请求和响应。

[0070] 根据一个实施例,在引导时,可以利用 $D\text{-MAC} = 01:00:5e:00:00:xx$ 、 $S\text{-MAC} = \text{NIC MAC 地址}$ 以及 $\text{EtherType} = 0xFFFF$ 或 $0XXXXX$ 来形成分组,其中 XX 或 xx 尚未被定义,但是可以由网络管理员确定并且在配置时间为主机操作系统给出。因此,这些数值可以是全部两个实体所知的任何数值。

[0071] 随后,在所连接的 NIC 端口上向 MAAS 发送出以太网分组(优选地是 MAAP 分组),其可以是交换设备的一部分。接下来,交换设备在该交换设备的其中一个端口上接收所述分组。相应地,MAAP 请求分组连同源端口信息被发送到本地 MAAS。此时,本地 MAAS 运行地址分配算法,并且确定对应于主机操作系统的 MAC 地址 H-MAC。

[0072] 在一种方法中,所述地址分配算法可以包括提取出在其上接收到 MAAP 请求的交换机端口号“p”,并且利用控制软件将该端口传递到 MAAS。随后对 MAC 地址分配表进行解析,并且确定第一可用(未分配)MAC 地址。该地址随后被添加到 MAAP 分组中,并且所述地址被标记为不可用(已分配)。

[0073] 接下来,MAAS 通过把 H-MAC 地址写入到 MAAP 分组中并且把目的地地址改变到原始 MAAP 请求的原始源地址而改动 MAAP 分组。所述分组被发送出源端口。发出请求的主机操作系统在预定等待时间(t_{wait})接收到响应 MAAP 分组,并且随后开始将 H-MAC 用作其所有未来通信的源 MAC 地址。

[0074] 如果 MAAP 响应没有在等待时间(t_{wait})内到达,则主机操作系统可以重试所述处理两次或更多次。在所述重试没有接收到响应之后,主机操作系统开始利用预先编程的 NIC MAC 地址作为其所有未来通信的源 MAC 地址。

[0075] 根据大多数实施例,这里所描述的方法和系统可以被初始化在平坦第 2 层网络上,其中所有的主机、设备或服务器都被指派相同子网的 IP 地址。

[0076] 现在参照图 6,其中示出了根据一个实施例的用于取回 MAC 地址的方法 600 的流程图。在各个实施例中,特别可以根据本发明在图 1-4 所描绘的任何环境中实施方法 600。当然,本领域技术人员通过阅读本发明的描述将会理解的是,在方法 600 中可以包括比在图 6 中具体描述的那些操作更多或更少的操作。

[0077] 方法 600 的每一个步骤可以由操作环境的任何适当组件实施。举例来说,在一个实施例中,方法 600 可以部分地或完全地由交换设备和/或交换设备的 MAAS 实施。

[0078] 如图 6 中所示,方法 600 可以发起于操作 602,其中向 MAAS 发送针对 MAC 地址的请求。在一种方法中,所述请求可以是 MAAP 分组。

[0079] 在操作 604 中,等待针对请求的响应。如果在预定时间量期间没有接收到针对所述请求的响应,则可以采取其他行动。

[0080] 在操作 606 中,当在预定时间量内接收到针对请求的响应时,使用 MAC 地址。所述响应包括 MAC 地址,并且在接收到 MAC 地址之后,由最初发送请求的设备或者耦合到该设备的某一其他设备使用所述 MAC 地址。

[0081] 在一个实施例中,当在预定时间量期间没有接收到针对请求的响应时,发送针对

MAC地址的请求的设备可以使用预先编程的MAC地址。根据一种方法,该MAC地址在制造时被预先编程到设备中。

[0082] 根据这里给出的实施例和方法,可以在不使用查找表的情况下做出交换决定,这是因为网络中的每一个设备具有对应于交换设备的目的地端口的MAC地址。此外,这样提供了极快的交换算法并且提供了非常低的等待时间。其还通过简化交换逻辑大大缩短了直通(cut-through)等待时间并且降低了成本。

[0083] 虽然前面描述了各个实施例,但是应当理解的是其仅仅是作为实例而非限制给出的。因此,本发明的实施例的宽度和范围不应当受限于任何前面所描述的示例性实施例,而是应当仅根据所附权利要求书及其等效表述来限定。

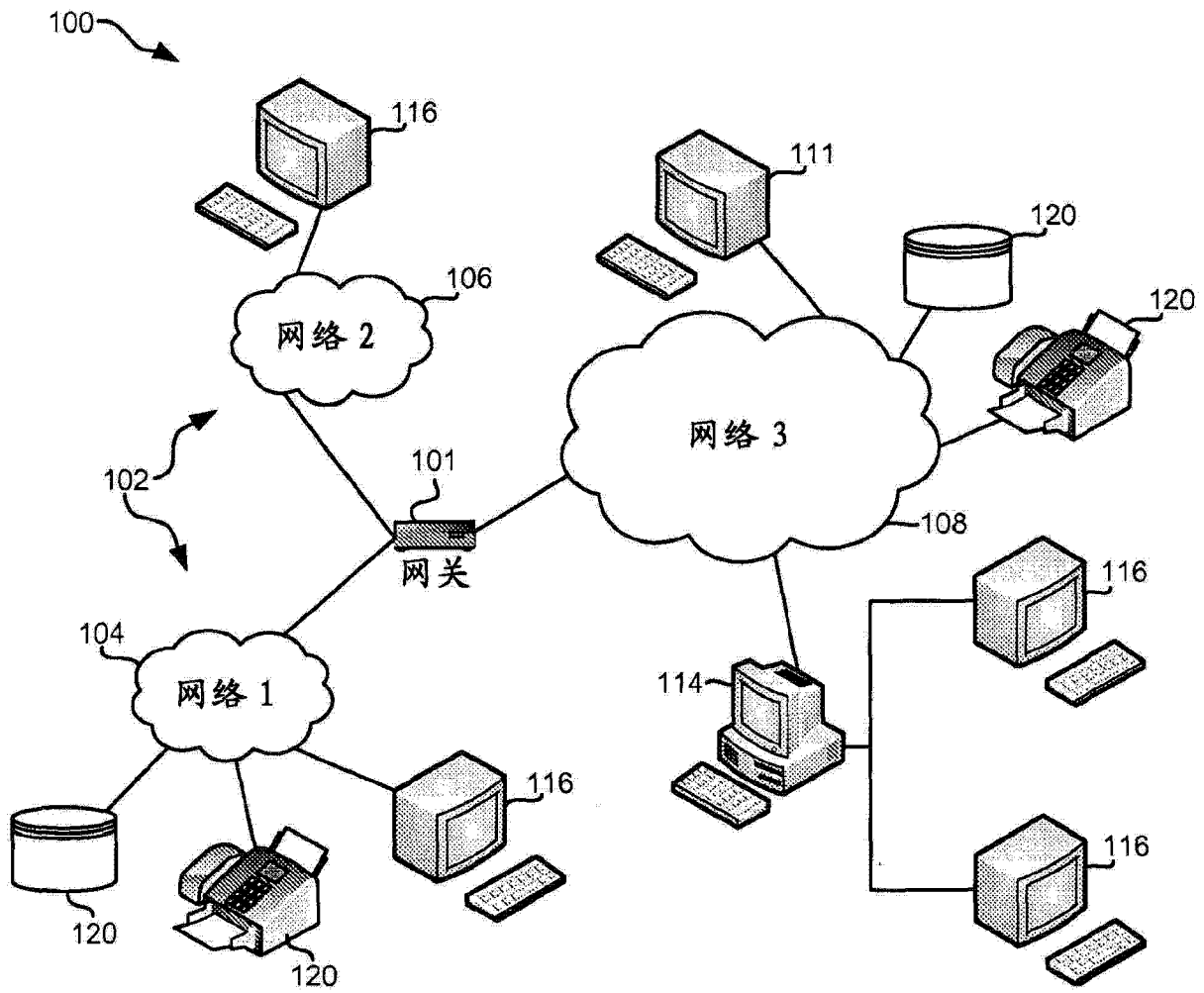


图 1

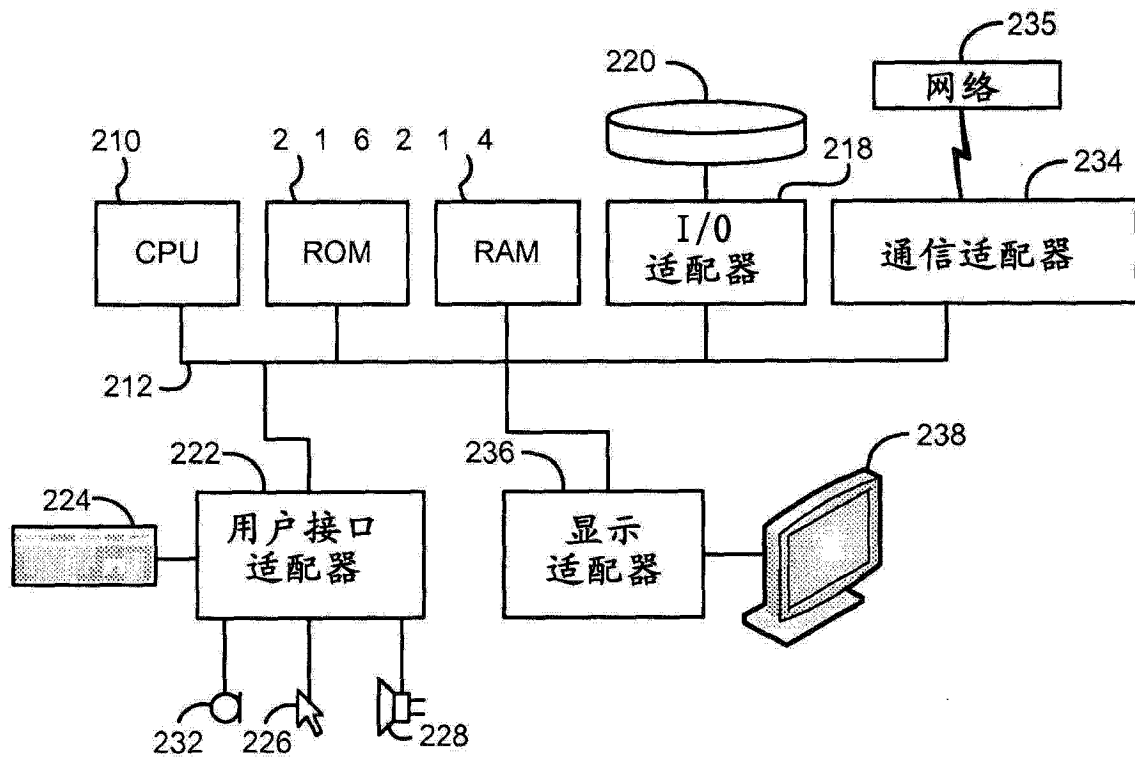


图 2

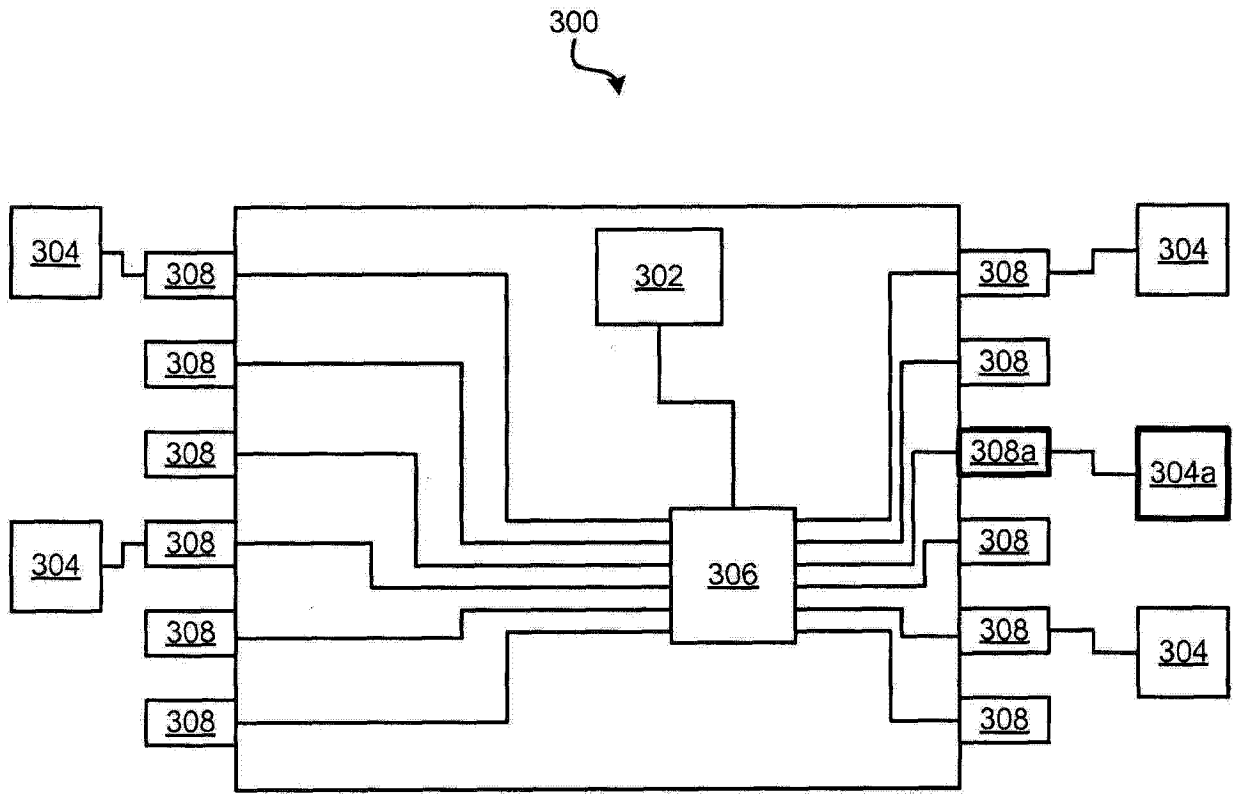


图 3

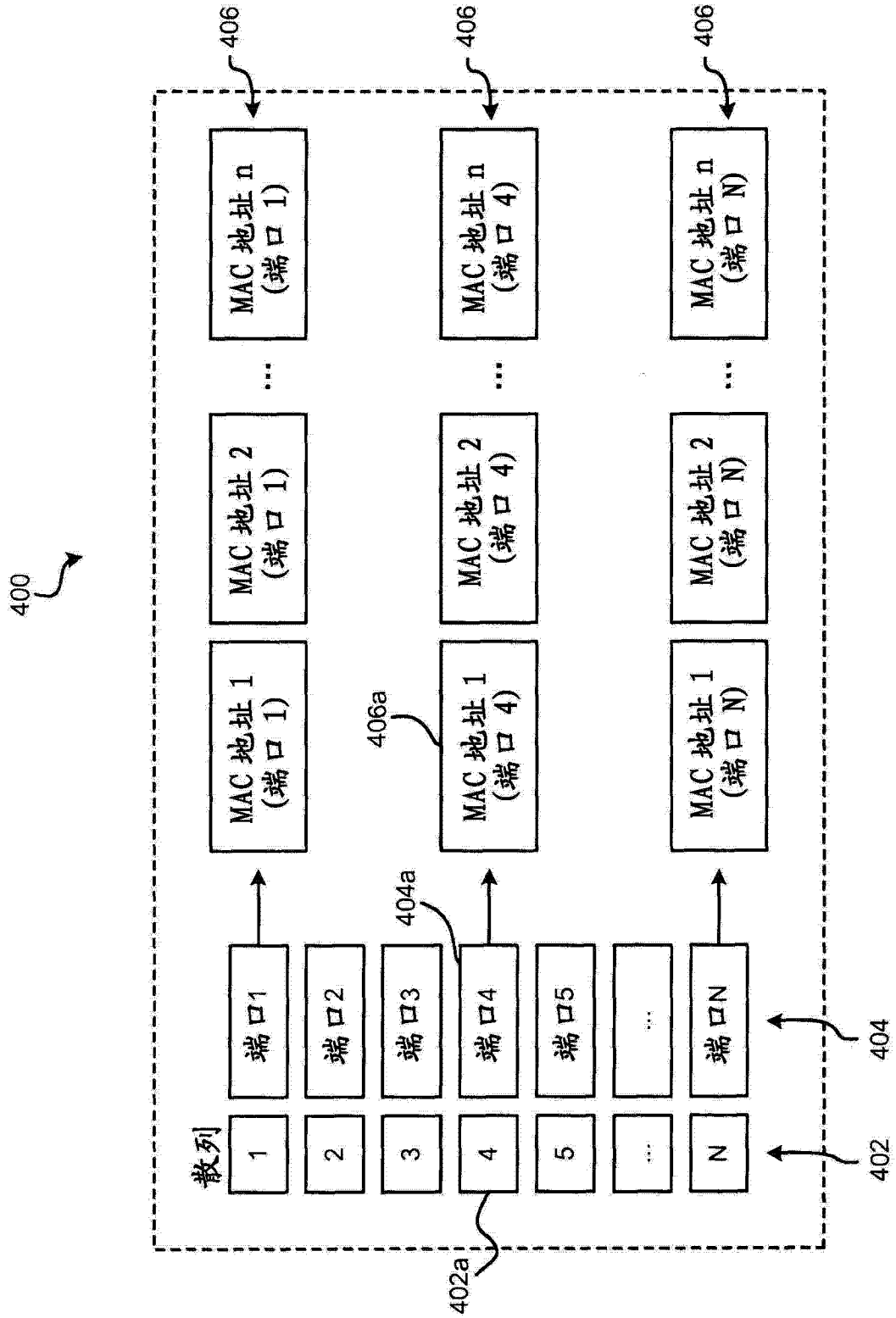


图 4

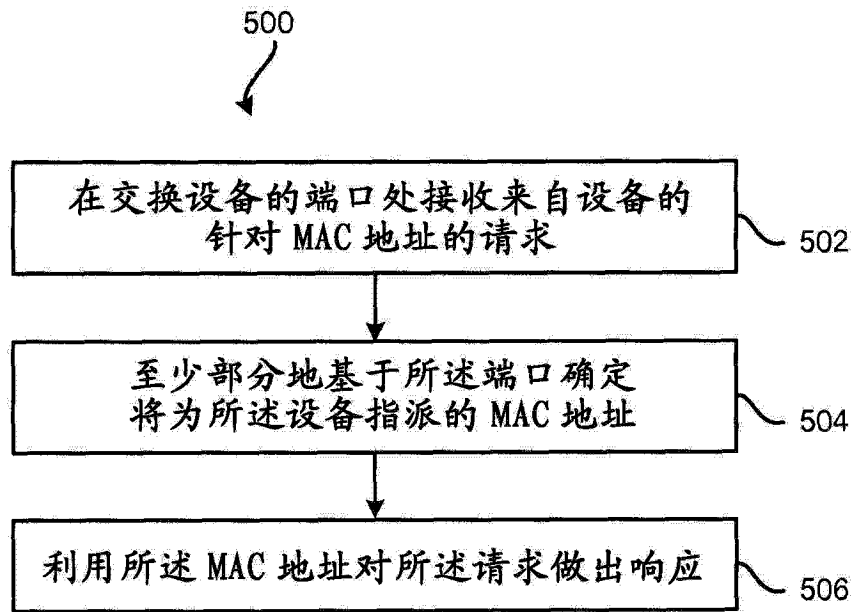


图 5

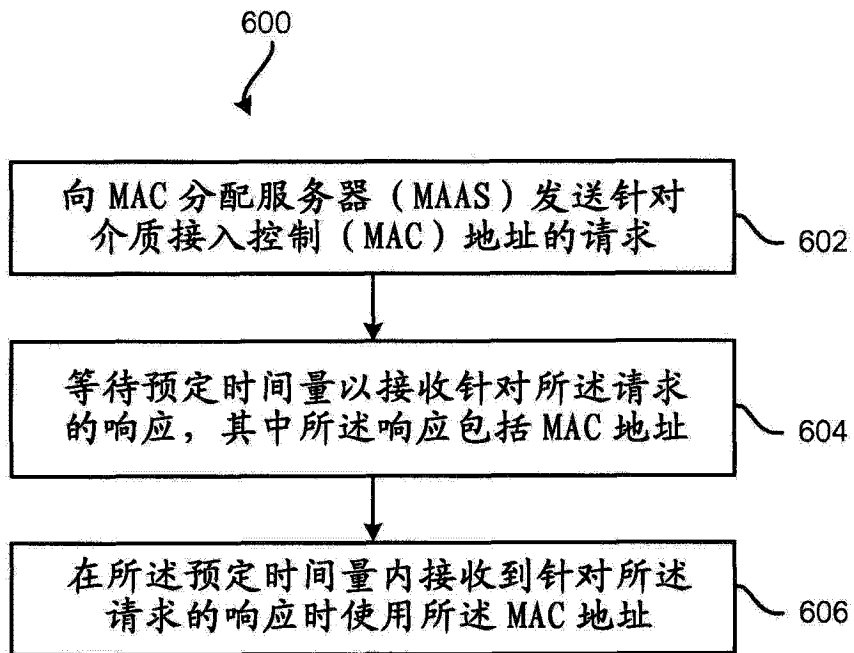


图 6