



# (12)发明专利申请

(10)申请公布号 CN 106656554 A

(43)申请公布日 2017. 05. 10

(21)申请号 201610896186.2

(22)申请日 2016.10.14

(71)申请人 盛科网络(苏州)有限公司

地址 215021 江苏省苏州市工业园区星汉街5号(腾飞工业坊)B幢4楼13/16单元

(72)发明人 丁奕

(74)专利代理机构 苏州集律知识产权代理事务所(普通合伙) 32269

代理人 安纪平

(51) Int. Cl.

H04L 12/24(2006.01)

H04L 12/709(2013.01)

H04L 12/911(2013.01)

权利要求书1页 说明书5页 附图4页

## (54)发明名称

MLAG环境下实现LACP的方法及装置

## (57)摘要

本发明揭示了一种在MLAG环境下实现LACP的方法及装置,方法包括如下步骤,在支持MLAG的设备之间进行LACP系统优先级同步,以保证两台支持MLAG的设备发出的LACPDU中的本端系统优先级的值相同;在支持MLAG的设备之间进行LACP系统ID同步,以保证两台支持MLAG的设备发出的LACPDU中的本端系统ID的值相同;在使能了MLAG的设备端口发送LACPDU时,使用端口上的MLAG ID替换端口聚合组ID填入本端操作Key中,并在本端操作Key中设置有能够防止误匹配的特殊标记。通过对传统LACP协议进一步的扩展,能够支持在MLAG环境下实现跨设备的聚合功能,且不需要借助其他协议。

101. 在支持MLAG的设备之间进行LACP系统优先级同步,以保证两台支持MLAG的设备发出的LACPDU中的本端系统优先级的值相同

202. 在支持MLAG的设备之间进行LACP系统ID同步,以保证两台支持MLAG的设备发出的LACPDU中的本端系统ID值相同

303. 在使能了MLAG的设备端口发送LACPDU时,使用端口上的MLAG ID替换端口聚合组ID填入本端操作Key中,且在本端操作Key中设置有能够防止误匹配的特殊标记

1. 一种在MLAG环境下实现LACP的方法,其特征在于,所述方法包括如下步骤:

S101,在支持MLAG的设备之间进行LACP系统优先级同步,以保证两台支持MLAG的设备发出的LACPDU中的本端系统优先级的值相同;

S202,在支持MLAG的设备之间进行LACP系统ID同步,以保证两台支持MLAG的设备发出的LACPDU中的本端系统ID值相同;

S303,在使能了MLAG的设备端口发送LACPDU时,使用端口上的MLAG ID替换端口聚合组ID填入本端操作Key中,且在本端操作Key中设置有能够防止误匹配的特殊标记。

2. 根据权利要求1所述的方法,其特征在于,所述LACP系统优先级配置成相同或不相同的值。

3. 根据权利要求1所述的方法,其特征在于,所述在支持MLAG设备之间进行LACP系统ID同步中,将LACP系统ID填入本端系统ID中。

4. 根据权利要求1所述的方法,其特征在于,所述特殊标记为将本端操作Key最高位的值置成1。

5. 根据权利要求4所述的方法,其特征在于,所述本端操作Key的值为0x8001,高字节0x80标识MLAG以区分传统LACP,低字节0x01标识MLAG ID的值为1。

6. 一种在MLAG环境下实现LACP的装置,其特征在于,包括LACP系统优先级同步模块,LACP系统ID同步模块,以及本端操作Key同步模块,其中,

LACP系统优先级同步模块在支持MLAG设备之间进行LACP系统优先级同步,

LACP系统ID同步模块在支持MLAG设备之间进行LACP系统ID同步,以及

本端操作Key同步模块使用端口上的MLAG ID替换端口聚合组ID,并填入本端操作Key中,且在本端操作Key中设置有能够防止误匹配的特殊标记。

7. 根据权利要求6所述的装置,其特征在于,所述LACP系统优先级同步模块中LACP系统优先级配置成相同或不相同的值。

8. 根据权利要求6所述的装置,其特征在于,所述LACP系统ID同步模块中将LACP系统ID填入本端系统ID中。

9. 根据权利要求6所述的装置,其特征在于,所述特殊标记为将本端操作Key最高位的值置成1。

10. 根据权利要求9所述的装置,其特征在于,所述本端操作Key的值为0x8001,高字节0x80标识MLAG以区分传统LACP,低字节0x01标识MLAG ID的值为1。

## MLAG环境下实现LACP的方法及装置

### 技术领域

[0001] 本发明涉及一种网络通信技术领域,尤其是涉及一种在MLAG环境下实现LACP的方法。

### 背景技术

[0002] 生成树协议(STP,Spanning Tree Protocol)是工作在OSI参考模型中二层的协议。在二层网络中,交换机起到了非常重要的作用,网络中,如果交换机出现故障,则会影响网络的使用,为了避免存在单点故障,在实际的二层链路上采用链路冗余,也就是采用交换设备之间多条链路连接,即多台交换机之间都有连接,这样即使一台交换机出现故障,也可以使用其他交换机,从而保证网络畅通。

[0003] 上述方式虽然能够保持网络畅通,但会产生一些问题。由于各个交换机设备之间都有链路连接,数据报到达目的主机的路线增多,从而使数据报在交换机之间不断的被转发,最终形成一个环路。环路容易导致广播风暴、帧重复复制、MAC地址表不稳定等问题,最终导致网络崩溃。

[0004] 生成树协议(STP,Spanning Tree Protocol)可以使负载均衡和冗余在二层实现。由于STP的存在,负载均衡不能在不同的设备之间实现,因为SPT会阻塞端口,从而防止网络存在环路,但带来的问题是会导致带宽的下降。通过跨设备链路聚合(MLAG,multi-chassis link aggregation)可以解决这个问题。在两台聚合交换机之间设置MLAG链路进行连接,使其在逻辑上如同一台设备,两台设备的端口共同形成聚合口,使得所有端口可以共同参与数据流量的转发。

[0005] 设备上一般都会使用聚合端口,设备上的聚合端口有两种形式,一种是静态聚合,一种是动态聚合。静态聚合顾名思义是静态配置的,由管理员静态指定聚合成员,不存在协议交互。动态聚合是基于IEEE802.3ad标准的LACP,是通过LACPDU(Link Aggression Control Protocol Data Unit)来进行信息交互的。传统的静态聚合支持MLAG没有问题,但使用动态LACP时,在MLAG场景下会存在端口不能绑定的问题。传统设备在使用LACP时,绑定在同一个聚合组里的端口都是存在同一个设备上的,但MLAG场景中,实现跨设备聚合。聚合端口会来自不同的设备,来自不同的设备的LACPDU就不能保证Actor Key一直,最终导致不能绑定在同一个聚合组之内。

### 发明内容

[0006] 本发明的目的在于克服现有技术的缺陷,提供一种在MLAG环境下实现LACP的方法,在交换机之间使用静态LACP或者动态LACP实现连接,不需要借助其他协议,从而解决传统的LACP在MLAG场景下不能正常工作的问题。

[0007] 为实现上述目的,本发明提出如下技术方案:一种在MLAG环境下实现LACP的方法,所述方法包括如下步骤:

[0008] S101,在支持MLAG的设备之间进行LACP系统优先级同步,以保证两台支持MLAG的

设备发出的LACPDU中的本端系统优先级的值相同；

[0009] S202,在支持MLAG的设备之间进行LACP系统ID同步,以保证两台支持MLAG的设备发出的LACPDU中的本端系统ID值相同；

[0010] S303,在使能了MLAG的设备端口发送LACPDU时,使用端口上的MLAG ID替换端口聚合组ID填入本端操作Key中,且在本端操作Key中设置有能够防止误匹配的特殊标记。

[0011] 优选地,所述LACP系统优先级配置成相同或不相同的值。

[0012] 优选地,所述在支持MLAG设备之间进行LACP系统ID同步中,将LACP系统ID填入本端系统ID中。

[0013] 优选地,所述特殊标记为将本端操作Key最高位的值置成1。

[0014] 优选地,所述本端操作Key的值为0x8001,高字节0x80标识MLAG以区分传统LACP,低字节0x01标识MLAG ID的值为1。

[0015] 一种在MLAG环境下实现LACP的装置,包括LACP系统优先级同步模块,LACP系统ID同步模块,以及本端操作Key同步模块,其中,

[0016] LACP系统优先级同步模块在支持MLAG设备之间进行LACP系统优先级同步,

[0017] LACP系统ID同步模块在支持MLAG设备之间进行LACP系统ID同步,以及

[0018] 本端操作Key同步模块使用端口上的MLAG ID替换端口聚合组ID,并填入本端操作Key中,且在本端操作Key中设置有能够防止误匹配的特殊标记。

[0019] 优选地,所述LACP系统优先级同步模块中LACP系统优先级配置成相同或不相同的值。

[0020] 优选地,所述LACP系统ID同步模块中将LACP系统ID填入本端系统ID中。

[0021] 优选地,所述LACP System ID同步模块中将System ID填入Actor System中。

[0022] 优选地,特殊标记为将本端操作Key最高位的值置成1。

[0023] 优选地,所述本端操作Key的值为0x8001,高字节0x80标识MLAG以区分传统LACP,低字节0x01标识MLAG ID的值为1。

[0024] 本发明的有益效果是:

[0025] 本发明所述的在MLAG环境下实现LACP的方法及装置,对传统LACP协议进一步的扩展,通过使用MLAG ID替换端口聚合组ID,使得在MLAG环境下能够实现跨设备聚合,且不需要借助其他协议,易于实现。

## 附图说明

[0026] 图1是本发明的实施例拓扑图示意图；

[0027] 图2是本发明的LACP报文结构示意图；

[0028] 图3是本发明的在MLAG环境下实现LACP的方法流程图示意图；

[0029] 图4是本发明的LACP报文详细信息示意图；

[0030] 图5是本发明的在MLAG环境下实现LACP的装置框图示意图。

[0031] 附图标记:

[0032] 1、LACP系统优先级同步模块,2、LACP系统ID同步模块,3、本端操作Key同步模块。

## 具体实施方式

[0033] 下面将结合本发明的附图,对本发明实施例的技术方案进行清楚、完整的描述。

[0034] IEEE802.3ad标准的LACP (Link Aggression Control Protocol,链路聚合控制协议) 是一个关于动态链路聚合的协议,它通过协议报文LACPDU (Link Aggression Control Protocol Data Unit,链路聚合控制协议数据单元) 和相连的设备交互信息。

[0035] 当端口启用LACP协议后,端口通过发送LACPDU来通告自己的系统优先级 (LACP System Priority)、设备MAC、端口的优先级、端口号和操作Key (Actor Key) 等。相连设备收到该报文后,根据所存储的其他端口的信息,选择端口进行相应的聚合操作,从而可以使双方在端口退出或者加入聚合组上达到一致。

[0036] 聚合组内的成员可能有3种状态,

[0037] (1) 当端口的链路状态处于Down时,端口不转发任何的数据报文,显示为“down”状态。

[0038] (2) 当端口的链路状态处于UP时,并经过LACP协商后,端口被置于聚合状态(端口被作为一个聚合组的一个成员参与聚合组的数据报文进行转发),显示为“bndl”状态;

[0039] (3) 当端口的链路状态处于UP时,但是由于对端没有启用LACP,或者因为端口属性和主端口不一致等一些因素导致经过协商端口被至于挂起状态(处于挂起状态的端口不参与数据报文转发),显示为“sups”状态。

[0040] 每台设备仅存在一个LACP聚合系统。每个LACP聚合系统都有唯一的系统优先级 (LACP System Priority)。本端系统与对端系统进行协商时,通常根据LACP的系统ID优先级来决定端口的转发状态,LACP的系统ID优先级较高的决定端口状态,其中,LACP的系统ID (LACP System ID) 由LACP的系统优先级和设备MAC地址组成,协商时,首先比较LACP的系统优先级,LACP的系统优先级越小,则LACP的系统ID的优先级越高,若在LACP的系统优先级相同的情况下,则进一步比较设备MAC地址,设备MAC地址越小,则LACP的系统ID的优先级越高。

[0041] 结合图1、图2和图3所示,本发明所揭示的一种在MLAG环境下实现LACP的方法,本实施例中,以三台交换机为例,对本发明所述的技术方案进行详细的说明。

[0042] 其中交换机1和交换机2为支持MLAG的设备、交换机3为不支持MLAG的设备,但支持LACP。如图2所示,LACP报文的详细信息,其中本端系统优先级 (Actor System Priority)、本端系统ID (Actor System) 和本端操作Key (Actor Key) 是根据本地/本端信息填充在报文中的。对端设备收到协议报文之后,上述三项会参与协议的运算。只有这三项信息完全相同的成员,才能被绑定成动态聚合口。在传统的LACP中,由于一侧端口都是在一台设备上的,因此本端系统优先级和本端系统ID能够保持一致,本端操作Key可以由管理员进行配置。比如系统将接口 (Interface) 上的端口聚合组ID (Channel-group ID) 作为本端操作Key (Actor Key) 填入LACP报文中,管理员可以通过配置相同的Channel-group ID来保证Actor Key的一致,因而在一台设备上的多个端口就可以成功被绑定到同一个聚合组里去了。

[0043] 如图1所示,但对于MLAG环境之下,交换机3上Channel-group 3中的两个端口分别来自两台不同的设备,即分别来自交换机1和交换机2,在两台不同设备上Channel-group ID也是很有可能不一致的,图中分别为Channel-group 1和Channel-group 2。这样他们分别发出的LACPDU中的本端系统优先级 (Actor System Priority)、本端系统ID (Actor System) 和本端操作Key (Actor Key) 可能都是不一样的,因此,在交换机3上就无法被绑定

到同一个聚合组之中。

[0044] 要实现MLAG环境下的LACP的正常功能,就需要保证上面三项内容的完全一致。通过三个步骤进行实现,如图3所示;

[0045] 在MLAG环境下实现LACP的方法,包括如下步骤:

[0046] S101,在支持MLAG设备之间进行LACP系统优先级(LACP System Priority)同步,以保证两台MLAG设备发出的LACPDU中的本端系统优先级(Actor System Priority)值相同;

[0047] S202,在支持MLAG设备之间进行LACP系统ID(LACP System ID)同步,以保证两台MLAG设备发出的LACPDU中的本端系统ID(Actor System)值相同;

[0048] S303,在使能了MLAG的设备端口发送LACPDU时,使用端口上的MLAG ID替换端口聚合组ID(Channel-group ID),并填入本端操作Key(Actor Key)中,且在本端操作Key(Actor Key)中设置有能够防止误匹配的特殊标记。

[0049] 更进一步的,在支持MLAG设备之间进行LACP系统优先级的同步中,由于LACP系统优先级在两台设备上可以分别配置,两台设备可以将优先级配成不同的值,也可以配置成相同的值。具体配置命令如下:Router(config)#lacp system-priority system-priority。由于支持MLAG设备有主从设备之分,从设备能够记录主设备发来的优先级。当使能了MLAG的设备端口上发送LACPDU时,将系统优先级填入本端系统优先级,以保证主从设备LACPDU中的本端系统优先级一致。

[0050] 更进一步的,在支持MLAG设备之间进行LACP系统ID的同步中,从设备记录主设备的LACP系统ID。在使能了MLAG的设备端口发送LACPDU的时,将LACP系统ID填入本端系统ID中,保证主从设备上LACPDU报文中的本端系统ID一致。

[0051] 更进一步的,由于两台设备绑定成MLAG口的聚合组,端口聚合组ID可以不同,因此使用MLAG ID来替换端口聚合组ID,能够保证本端操作Key一致,本端操作Key是管理员可以保证的配置。因此需要扩展传统的LACP协议来实现端口聚合,通过使用MLAG ID来替换端口聚合组ID填入本端操作Key之中。

[0052] 为了和没有使能了MLAG的设备端口发出的LACPDU区分开,通过将本端操作Key的进行特殊标记避免误匹配,本实施例中,将本端操作Key的最高位(Bit)置成1,当然也可以选用其他方式进行标记,如将次高位置为1,目的是为了能够有效地与本地LACP进行区分开。如交换机1上的端口使能了MLAG,并且MLAG ID配置成1,交换机2上的端口没有使能MLAG,只是把端口加入了端口聚合1,这时候如果没有最高bit的区别,那么两者的本端操作Key将都是0x1,在交换机3上做聚合处理的时,能够将两个原本不该绑定在一个聚合组的端口进行误绑定。综上所述,交换机2发出的LACPDU的本端操作Key为0x1,交换机1发出的本端操作Key为0x8001,因为Actor Key字段为16位(Bit),即两个字节(Byte),由于两者Actor Key不同,从而避免被误绑定在同一个聚合组。另外,不考虑聚合组ID超过32768的情况,即超过0x8000,通常系统中不会支持这么多聚合组。

[0053] 如图4所示,从网络系统中抓包可以看出,Actor Key已经被修改成0x8001了。高字节的0x80是标识MLAG以区分传统LACP的,低字节的0x01是代表MLAG ID是1。经过这样的处理之后,交换机3收到的分别来自交换机1和交换机2的两台设备的LACPDU中的本端系统优先级、本端系统ID和本端操作Key三项内容是一致的,因此在交换机3上这两个端口就可以

成功被绑定到同一个聚合组之中。更重要的是,本发明只需要被实施应用在支持MLAG的设备之上如图所示的交换机1和交换机2上,并不需要被实施在不支持MLAG的普通交换机设备上,如图中所示的交换机3。对于不支持MLAG的普通设备,只支持传统的LACP协议就能够实现端口聚合。

[0054] 更进一步的,如图5所示,一种在MLAG环境下实现LACP的装置,包括LACP系统优先级同步模块,LACP系统ID同步模块,以及本端操作Key同步模块,其中,

[0055] LACP系统优先级同步模块在支持MLAG设备之间进行LACP系统优先级同步,

[0056] LACP系统ID同步模块在支持MLAG设备之间进行LACP系统ID同步,以及

[0057] 本端操作Key同步模块使用端口上的MLAG ID替换端口聚合组ID,并填入本端操作Key中,且在本端操作Key (Actor Key) 中设置有能够防止误匹配的特殊标记。

[0058] 具体的,在LACP系统优先级同步模块中使用Router(config)#lacp system-priority system-priority将两台设备的优先级配成不同的值,也可以配置成相同的值,当使能了MLAG的设备端口上发送LACPDU时,将优先级填入本端系统优先级,以保证主从设备LACPDU中的本端系统优先级一致。

[0059] 在LACP系统ID同步模块中,将LACP系统ID填入本端系统ID中,保证主从设备上LACPDU报文中的本端系统ID一致。

[0060] 在本端操作Key同步模块中,使用端口上的MLAG ID替换端口聚合组ID,并填入本端操作Key中,并将本端操作Key进行特殊标记,避免误匹配,本实施例中,将本端操作Key的最高位(Bit)置成1,当然也可以选用其他方式进行标记,如将次高位置为1,目的是为了能够有效地与本地LACP进行区分开,具体的,所述本端操作Key的值为0x8001,高字节0x80标识MLAG以区分传统LACP,低字节0x01标识MLAG ID的值为1。

[0061] 本发明所述的MLAG环境下实现LACP的方法及装置,通过对传统LACP协议进一步的扩展,能够支持其在MLAG环境下实现跨设备的聚合功能,且不需要借助其他协议。

[0062] 本发明的技术内容及技术特征已揭示如上,然而熟悉本领域的技术人员仍可能基于本发明的教示及揭示而作种种不背离本发明精神的替换及修饰,因此,本发明保护范围应不限于实施例所揭示的内容,而应包括各种不背离本发明的替换及修饰,并为本专利申请权利要求所涵盖。

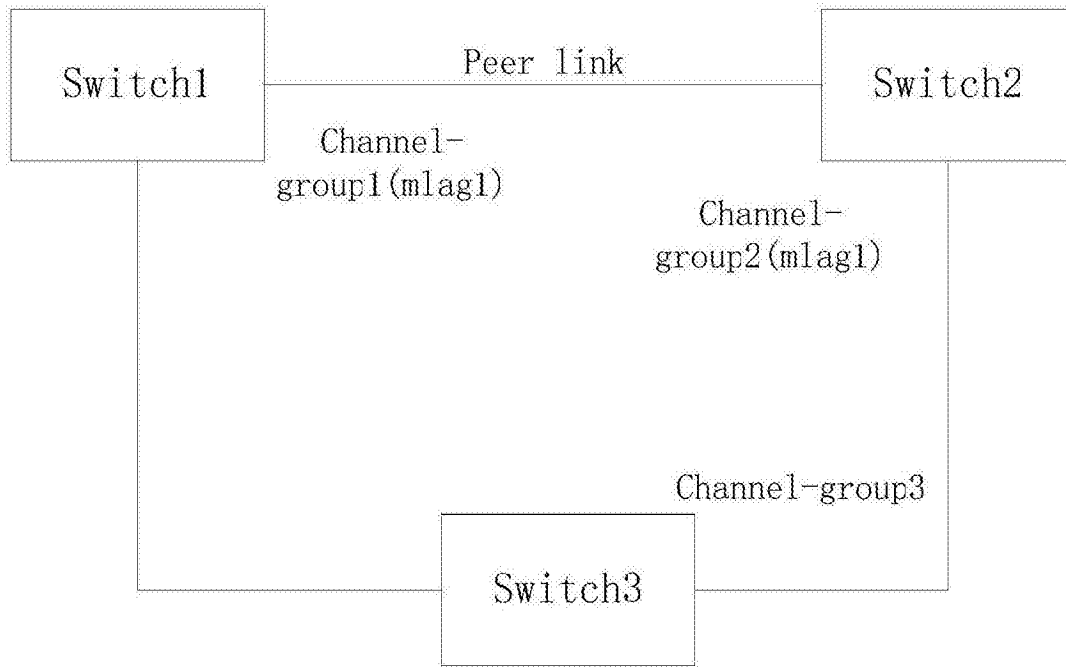


图1



Octets	
Destination Address	6
Source Address	6
Length/Type	2
Subtype = LACP	1
Version Number	1
TLV_type = Actor Information	1
Actor_Information_Length = 20	1
Actor_System_Priority	2
Actor_System	6
Actor_Key	2
Actor_Port_Priority	2
Actor_Port	2
Actor_State	1
3 Reserved	3
TLV_type = Partner Information	1
Partner_Information_Length = 20	1
Partner_System_Priority	2
Partner_System	6
Partner_Key	2
Partner_Port_Priority	2
Partner_Port	2
Partner_State	1
3 Reserved	3
TLV_type = Collector Information	1
Collector_Information_Length = 16	1
CollectorMaxDelay	2
Reserved	12
TLV_type = Terminator	1
Terminator_Length = 0	1
Reserved	50
FCS	4

图2

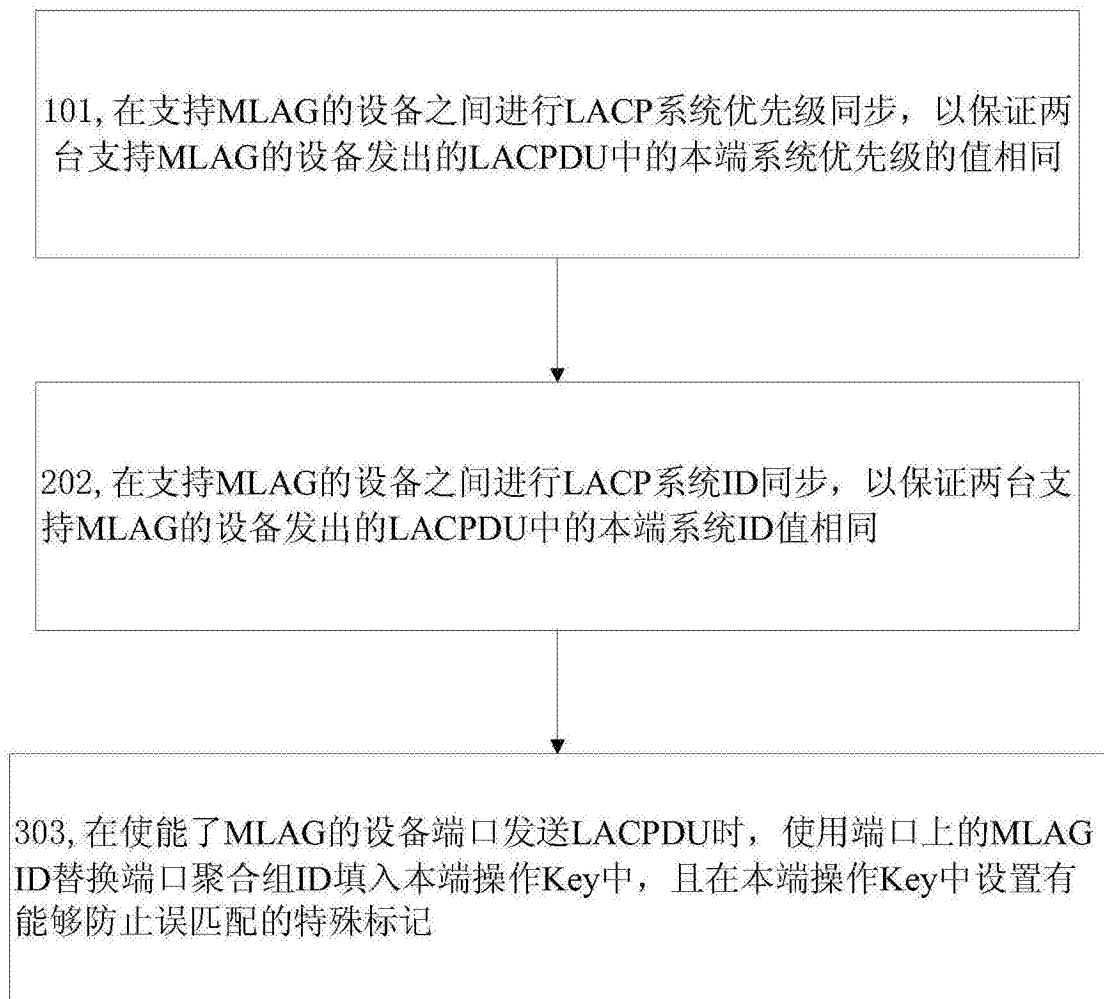


图3

```

0 0.603188      32:1d:b6:6b:4d:0d      Slow-Protocols      LACP      124 Link Aggregation
0 0.836838      0e:a2:bc:49:75:0d      Slow-Protocols      LACP      124 Link Aggregation

# Frame 0: 124 bytes on wire (992 bits), 124 bytes captured (992 bits)
# Ethernet II, Src: 32:1d:b6:6b:4d:0d (32:1d:b6:6b:4d:0d), Dst: Slow-Protocols (01:80:c2:00:00:02)
# Link Aggregation Control Protocol
  Slow Protocols subtype: LACP (0x01)
  LACP version Number: 0x01
  Actor Information: 0x01
  Actor Information Length: 0x14
  Actor System Priority: 32768
  Actor System: 3a:ff:a2:00:2d:00 (3a:ff:a2:00:2d:00)
  Actor Port Priority: 32768
  Actor Port: 13
# Actor State: 0x0d (Activity, Aggregation, Synchronization)
  Reserved: 000000
  Partner Information: 0x02
  Partner Information Length: 0x14
  Partner System Priority: 32768
  Partner System: 0e:a2:bc:49:75:00 (0e:a2:bc:49:75:00)
  Partner Key: 1
  Partner Port Priority: 32768
  Partner Port: 13
# Partner State: 0x15 (Activity, Aggregation, Collecting)
  Reserved: 000000
  Collector Information: 0x03
  Collector Information Length: 0x10
  collector Max Delay: 5
  Reserved: 00000000000000000000000000000000
  Terminator Information: 0x00
  Terminator Length: 0x00
  Reserved: 0000000000000000000000000000000000000000000000000000000000000000...

0000  01 80 c2 00 00 02 32 1d b6 6b 4d 0d 88 09 01 01  .....2...KM.....
0010  01 14 80 00 3a ff a2 00 2d 00 80 00 00 00 00 00 00  .....
0020  0d 00 00 00 02 14 80 00 0e e2 bc 49 75 00 00 01  .....TU.....
0030  80 00 00 0d 15 00 00 00 03 10 00 05 00 00 00 00  .....
0040  00 00 00 00 00 00 00 00 60 00 00 00 00 00 00 00  .....
0050  00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00  .....
0060  00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00  .....
0070  00 00 00 00 00 00 00 00 00 00 00 00  .....

```

图4

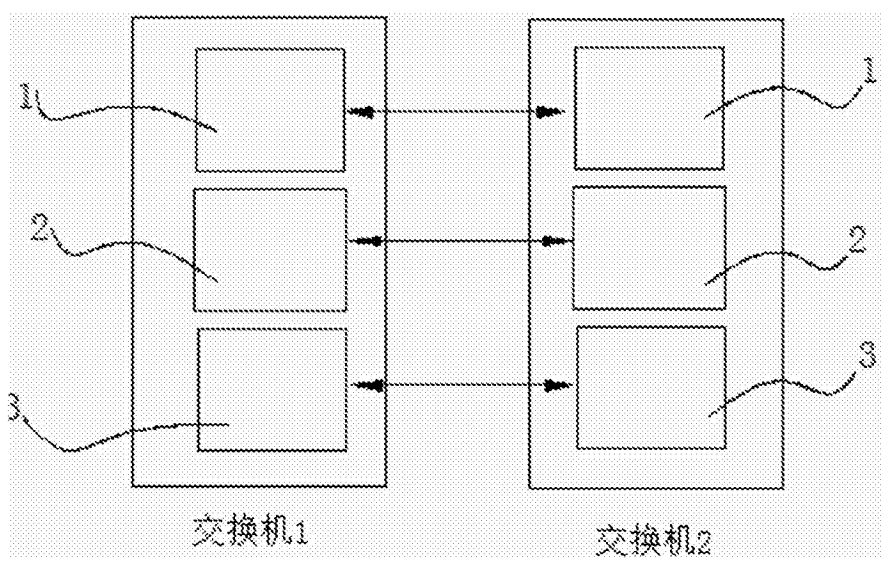


图5