



(12)发明专利

(10)授权公告号 CN 106656789 B

(45)授权公告日 2020.04.14

(21)申请号 201611270597.7

H04L 12/933(2013.01)

(22)申请日 2016.12.30

H04L 12/935(2013.01)

(65)同一申请的已公布的文献号
申请公布号 CN 106656789 A

(56)对比文件

US 2011228771 A1,2011.09.22,

(43)申请公布日 2017.05.10

审查员 寇利敏

(73)专利权人 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 龚海东 方沛昱 杨曙军

(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 12/709(2013.01)

H04L 12/931(2013.01)

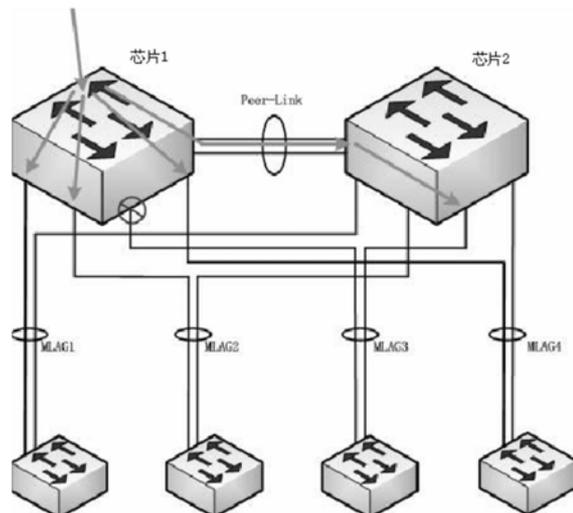
权利要求书1页 说明书4页 附图2页

(54)发明名称

MLAG广播和组播的芯片实现方法

(57)摘要

本发明揭示了一种MLAG广播和组播的芯片实现方法,通过将Peer-Link加入广播复制列表和组播的复制列表,把报文通过复制给MLAG对端设备,并使用端口隔离机制实现MLAG非预期端口的报文丢弃,在使用APS实现MLAG单播的基础上,实现了MLAG的广播和组播转发。



1. 一种MLAG广播和组播的芯片实现方法,其特征在于,所述方法包括:

S1,配置芯片:

将Peer-Link端口和VLAN成员端口加入复制组A,将Peer-Link端口和二层组播组成员端口加入复制组B,在Peer-Link端口入方向上设置源隔离ID为S1,在MLAG口出方向上设置目的隔离ID为D1,及配置隔离条目表;

S2,报文处理:

S21,报文进入MLAG本地芯片查转发表,得到复制组A/复制组B,查找复制组A/复制组B,获得复制组A/复制组B成员列表,复制报文到与MLAG本地芯片相连通的每一个MLAG成员;

S22,报文从Peer-Link进入与MLAG本地芯片相连的MLAG远端芯片后,查转发表得到复制组A/复制组B,查找复制表获得复制组A/复制组B成员列表,复制报文到第二芯片的每一个MLAG成员端口;

S23,所述第二芯片除Peer-Link端口外的每一个MLAG成员端口,使用源隔离ID查找隔离条目表,得到各自的目的隔离ID,根据目的隔离ID丢弃或发送报文出去。

2. 根据权利要求1所述的MLAG广播和组播的芯片实现方法,其特征在于,步骤S21或S22中,报文从MLAG本地芯片或MLAG远端芯片进入后,解析报文二层头,然后使用MACDA加VLAN ID查找二层转发表,得到复制组A/复制组B。

3. 根据权利要求1或2所述的MLAG广播和组播的芯片实现方法,其特征在于,所述报文为广播、未知单播、未知组播和已知组播中的一种,所述报文为广播、未知单播或未知组播时,查转发表,得到复制组A;所述报文为已知组播时,查转发表,得到复制组B。

4. 根据权利要求1所述的MLAG广播和组播的芯片实现方法,其特征在于,报文复制过程中,源地址和目的地址相同的报会自动丢弃。

5. 根据权利要求1所述的MLAG广播和组播的芯片实现方法,其特征在于,对于MLAG口出方向上down的端口,将目的隔离ID配置为0。

6. 根据权利要求5所述的MLAG广播和组播的芯片实现方法,其特征在于,所述目的隔离ID配置为0时,将复制的报文从MLAG口发送出去。

7. 根据权利要求1所述的MLAG广播和组播的芯片实现方法,其特征在于,隔离条目表中,目的隔离ID用bitmap标识。

8. 根据权利要求7所述的MLAG广播和组播的芯片实现方法,其特征在于,步骤S23中,根据查询到的第S1条条目的bitmap的从低位开始的第D1位置的值,对报文进行丢弃或发送处理。

9. 根据权利要求8所述的MLAG广播和组播的芯片实现方法,其特征在于,配置所述bitmap的从低位开始的第D1位为1时,丢弃复制出来的报文。

MLAG广播和组播的芯片实现方法

技术领域

[0001] 本发明涉及一种链路聚合技术,尤其是涉及一种MLAG广播和组播的芯片实现方法。

背景技术

[0002] 在高可靠性的数据中心拓扑中,典型的会通过两台聚合交换机来连接TOR (Top of Rack, 架顶式) 交换机和服务器的提供冗余保护。在这样的拓扑结构中,生成树协议 (Spanning-Tree Protocol, STP) 通过block (阻塞) 聚合交换机的一半的端口来防止网络环路,但这样做会降低50%的带宽。

[0003] 通过部署MLAG (多链路聚合) 可以解决这个问题。在两台聚合交换机的中间通过一条MLAG链路进行连接,使其在逻辑上如同一台设备。两台设备上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。

[0004] 现有,一般通过二层转发表的表项刷新实现MLAG,当MLAG口down (关闭) 的时候,把所有出口为MLAG出口的二层转发条目刷新成出口为peer link (同行的链接),当MLAG up (打开) 时,再从peer link口刷回MLAG口。

[0005] 但是上述技术会存在以下缺陷:1、速度比较慢,会造成较长时间的丢包和断流。2、MLAG up时,再从peer link口刷回MLAG口的情况,需要保存大量软表。

[0006] 现有也有通过APS (Automatic Protection Switching, 自动保护倒换) 机制实现MLAG单播。

[0007] 步骤1,创建好MLAG的peer link;

[0008] 步骤2,进行聚合端口和MLAG组的绑定。

[0009] 第一步和第二步的顺序并不重要,但在第三步进行之前,前面两步需要都完成。

[0010] 根据绑定了MLAG组的聚合端口的信息和peer link的信息,创建aps group (自动保护切换组)。芯片中的DsApsBridge这张表中的workingDestMap使用被绑定的聚合端口,protectingDestMap使用peer link所对应的端口或者聚合端口。创建完成这个保护组之后,能得到一个NEXTHOP ID,并且在软件层面保存MLAG ID和对应的NEXTHOP ID的对应关系。

[0011] 在进行FDB学习的时候,根据源端口属于哪个MLAG组,查找上述的对应关系,得到保护组对应的NEXTHOP ID,使用这个NEXTHOP ID进行FDB的添加。

[0012] 当检测到本地MLAG端口由up变down的时候,即当保护切换发生时,可以直接进行APS的切换,将DsApsBridge中的protectingEn置起来即可。

[0013] 当检测到本地MLAG端口由down变up的时候,即当保护恢复时,则直接进行APS的恢复,将DsApsBridge中的protectingEn清除掉即可。

[0014] 在聚合端口和MLAG组解除绑定关系,或者删除peer link,两者任何一个事件发生时,删除保护组,释放APS资源。

[0015] 但是,上述第二种技术中,APS机制不能用于广播和组播。

发明内容

[0016] 本发明的目的在于克服现有技术的缺陷,提供一种MLAG广播和组播的芯片实现方法,通过端口隔离的方式以实现MLAG的广播和组播转发。

[0017] 为实现上述目的,本发明提出如下技术方案:一种MLAG广播和组播的芯片实现方法,所述方法包括:

[0018] S1,配置芯片:

[0019] 将Peer-Link端口和VLAN成员端口加入复制组A,将Peer-Link端口和二层组播组成员端口加入复制组B,在Peer-Link端口入方向上设置源隔离ID为S1,在MLAG口出方向上设置目的隔离ID为D1,及配置隔离条目表;

[0020] S2,报文处理:

[0021] S21,报文进入MLAG本地芯片查转发表,得到复制组A/复制组B,查找复制组A/B,获得复制组A/B成员列表,复制报文到与MLAG本地芯片相连通的每一个MLAG成员;

[0022] S22,报文从Peer-Link进入与MLAG本地芯片相连的MLAG远端芯片后,查转发表得到复制组A/复制组B,查找复制表获得复制组A/B成员列表,复制报文到第二芯片的每一个MLAG成员端口;

[0023] S23,所述第二芯片除Peer-Link端口外的每一个MLAG成员端口,使用源隔离ID查找隔离条目表,得到各自的目的隔离ID,根据目的隔离ID丢弃或发送报文出去。

[0024] 优选地,步骤S21或S22中,报文从MLAG本地芯片或MLAG远端芯片进入后,解析报文二层头,然后使用MACDA加VLAN ID查找二层转发表,得到复制组A/复制组B。

[0025] 优选地,所述报文为广播、未知单播、未知组播和已知组播中的一种,所述报文为广播、未知单播或未知组播时,查转发表,得到复制组A;所述报文为已知组播时,查转发表,得到复制组B。

[0026] 优选地,所述S21或S22中,按复制组ID查找复制组A/B,获得复制组A/B成员列表。

[0027] 优选地,报文复制过程中,源地址和目的地址相同的报会自动丢弃。

[0028] 优选地,对于MLAG口出方向上down的端口,将目的隔离ID配置为0。

[0029] 优选地,所述目的隔离ID配置为0时,将复制的报文从MLAG口发送出去。

[0030] 优选地,隔离条目表中,目的隔离ID用bitmap标识。

[0031] 优选地,步骤S23中,根据查询到的第S1条条目的bitmap的从低位开始的第D1位置的值,对报文进行丢弃或发送处理。

[0032] 优选地,配置所述bitmap的从低位开始的第D1位为1时,丢弃复制出来的报文。

[0033] 与现有技术相比,本发明通过将Peer-Link加入广播复制列表和组播的复制列表,把报文通过复制给MLAG对端设备,并使用端口隔离机制实现MLAG非预期端口的报文丢弃,在使用APS实现MLAG单播的基础上,实现了MLAG的广播和组播转发。

附图说明

[0034] 图1是本发明MLAG广播和组播的芯片实现方法的流程示意图;

[0035] 图2是本发明报文处理的流程示意图;

[0036] 图3是本发明实施例的原理示意图。

具体实施方式

[0037] 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。

[0038] 本发明所揭示的一种MLAG广播和组播的芯片实现方法,在使用APS实现MLAG单播的基础上,通过端口隔离的方式实现了MLAG的广播和组播转发。

[0039] 结合图1~图3所示,本发明所揭示的一种MLAG广播和组播的芯片实现方法,包括以下步骤:

[0040] 步骤S1,配置芯片:

[0041] 将Peer-Link端口和VLAN成员端口加入复制组A,将Peer-Link端口和二层组播组成员端口加入复制组B,在Peer-Link端口入方向上设置源隔离ID为S1,在MLAG口出方向上设置目的隔离ID为D1,及配置隔离条目表。

[0042] 具体地,首先介绍下与本方案相关的通用机制:

[0043] VLAN内广播机制,把VLAN内的所有成员端口加入到一个复制组A内,在二层转发表(使用MACDA+VLAN ID进行查找)内配置以下条目:

[0044] MAC (FFFF.FFFF.FFFF)+VLAN ID→复制组A

[0045] VLAN默认条目→复制组A

[0046] 即将VLAN内的所有成员端口加入到复制组A内,查找二层转发表时,使用MACDA+VLAN ID进行查找,得到复制组A。

[0047] 二层组播机制,把二层组播组所有成员端口加入到一个复制组B内,在二层转发表内配置以下条目:

[0048] MAC (组播)+VLAN ID→复制组B

[0049] 即将组播组所有成员端口加入到一个复制组B内,查找二层转发表时,使用MAC (组播)+VLAN ID进行查找,得到复制组B。

[0050] 报文复制机制,按复制组ID查找复制组,获得复制组成员列表,然后复制报文到每一个成员,其中源和目的地相同的报会自动丢弃。

[0051] 与上述通用机制相比,本发明实施例的芯片配置为:

[0052] 将Peer-Link端口和VLAN成员端口加入复制组A。

[0053] 将Peer-Link端口和二层组播组成员端口加入复制组B。

[0054] 在Peer-Link入方向上设置源隔离ID为S1。

[0055] 在MLAG口出方向上设置目的隔离ID为D1,凡是MLAG口出方向上down的MLAG口,将隔离ID配置为0。

[0056] 配置隔离条目表,隔离条目表使用源隔离ID进行查找,得到一个标识目的隔离ID目的bitmap,对于down掉的MLAG口,将隔离条目表中第S1个条目的bitmap的从低位开始的第D1位置1。

[0057] 图3中所示的芯片1和芯片2均进行以上配置。

[0058] 步骤S2,报文处理:

[0059] S21,报文进入MLAG本地芯片查转发表,得到复制组A/复制组B,查找复制组A/B,获得复制组A/B成员列表,复制报文到MLAG本地芯片的每一个MLAG成员端口。

[0060] 具体地,以图3所示为例,芯片1为组成MLAG的MLAG本地芯片,芯片2为MLAG远端芯片。报文从源端口进入芯片1中,解析报文二层头,然后使用MACDA+VLAN ID查找二层转发

表,广播、未知单播和未知组播报文会查找到出口为复制组A;已知组播报文会查找到出口为复制组B;已知单播报文按转发表直接得到物理出口,不在本专利保护范围内,这里不做描述。

[0061] 然后,按照复制组ID查找复制组表,获得复制组成员列表,如查找到为复制组A,则按照复制组A的ID查找复制组A,获得复制组A的所有成员列表;查找到复制组B与复制组A的复制过程相同。

[0062] 然后复制报文到与芯片1相连通的每一个MLAG成员。如图3,芯片1的MLAG 1、2、4端口会直接复制报文到目标设备(设备1、设备2和设备4),同时芯片1的Peer-Link端口上也会复制一份报文给组成MLAG的另一台设备(即芯片2)。因芯片1的MLAG3端口down了,根据APS机制,流量会经过Peer-Link发往右侧设备(即芯片2),所以,芯片1的MLAG3端口复制不了报文到目标设备3。

[0063] 也就是说,从芯片1进入的报文,通过查找二层转发表,分别发送给了目标设备1、2和4,发往目标设备3的报文会经过Peer-Link发往芯片2。

[0064] S22,报文从Peer-Link进入与MLAG本地芯片相连的MLAG远端芯片后,查转发表得到复制组A/复制组B,查找复制表获得复制组A/B成员列表,复制报文到第二芯片的每一个MLAG成员端口。

[0065] 具体地,芯片2的Peer-Link端口在入方向收到报文,会从上述步骤S1的端口配置中得到源隔离ID,芯片2收到报文后,与芯片1中报文解析、查表、复制过程相同:解析报文二层头,然后使用MACDA+VLAN ID查找二层转发表,同样会得到复制组A/复制组B,按复制组ID查找复制组,获得复制组成员列表,然后复制报文到芯片2的每一个MLAG成员端口。芯片2的MLAG成员端口包含Peer-Link端口和MLAG1、2、3、4口。其中Peer-Link口由于源和目的地相同,复制的报文直接丢弃,不会再发回去。

[0066] 步骤S23,所述第二芯片除Peer-Link端口外的每一个MLAG成员端口,使用源隔离ID查找隔离条目表,得到各自的目的隔离ID,根据目的隔离ID丢弃或发送报文出去。

[0067] MLAG1、2、4口会在出方向上得到目的隔离ID为D1,使用源隔离ID S1进行隔离条目表查找,得到第S1条条目的bitmap,发现其中的地D1个bit为1,丢弃复制出来的报文;MLAG 3口由于目的隔离ID为0,不会丢弃报文,报文会正常从3口出去。

[0068] 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。

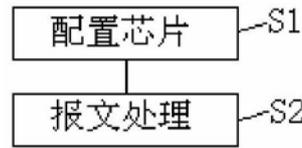


图1

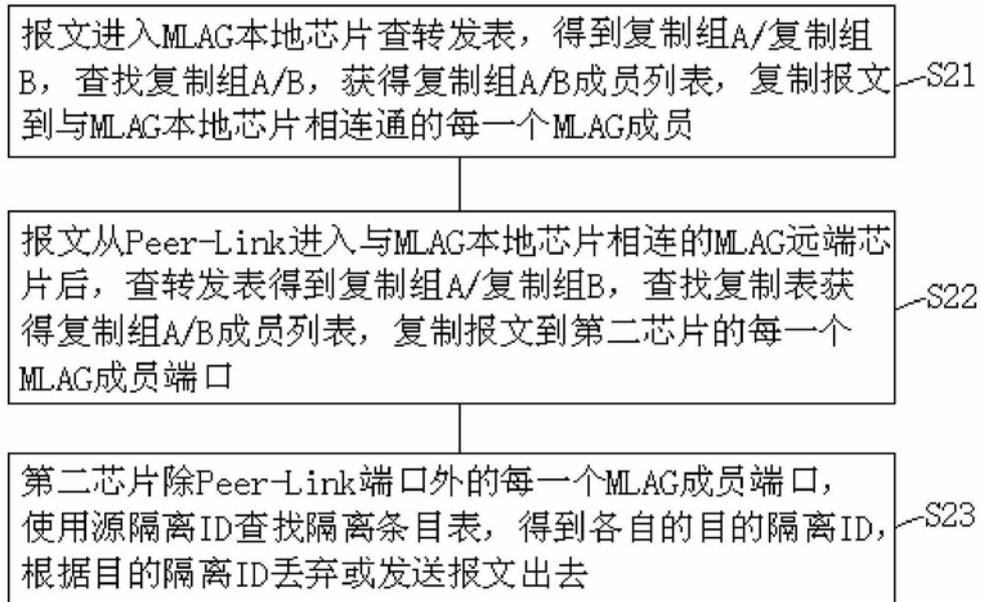


图2

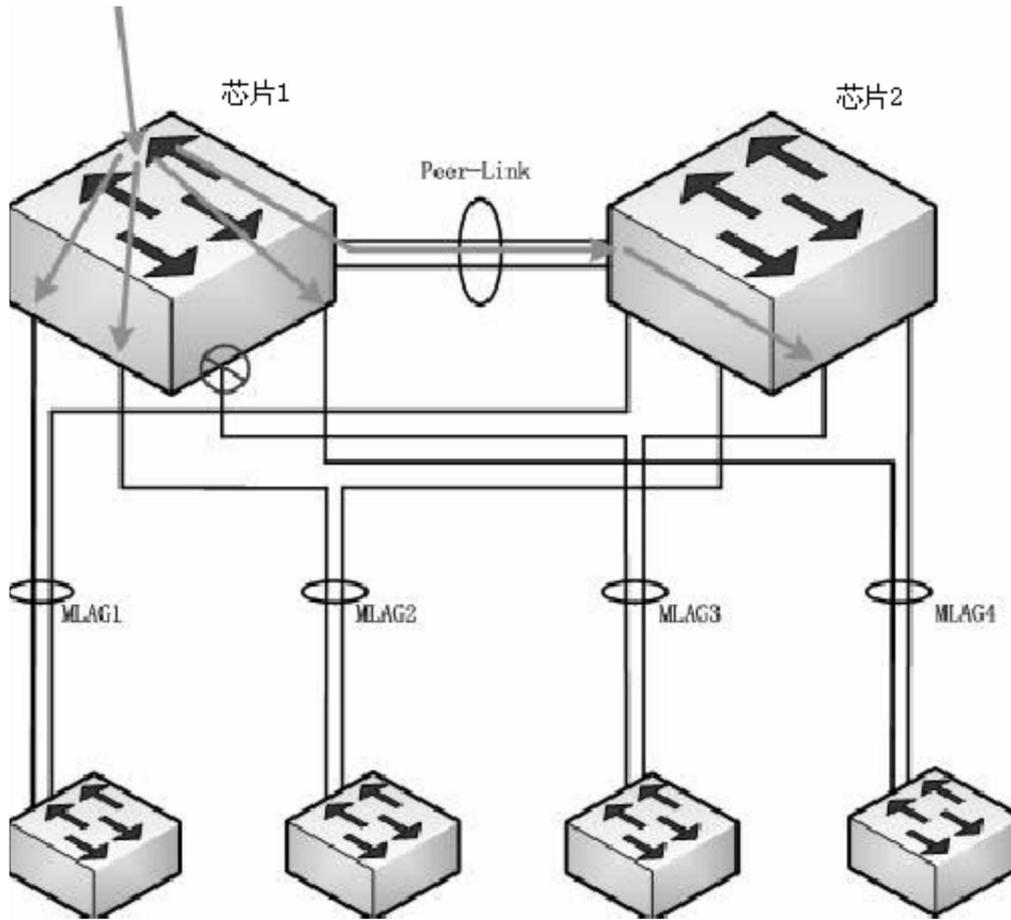


图3