



# (12) 发明专利申请

(10) 申请公布号 CN 104348637 A

(43) 申请公布日 2015. 02. 11

(21) 申请号 201310321034. 6

(22) 申请日 2013. 07. 26

(71) 申请人 中国科学院声学研究所

地址 100190 北京市海淀区北四环西路 21 号

申请人 北京中科海力技术有限公司

(72) 发明人 刘磊 郑艳伟 彭飞

(74) 专利代理机构 北京法思腾知识产权代理有限公司 11318

代理人 杨小蓉 杨青

(51) Int. Cl.

H04L 12/24 (2006. 01)

H04L 29/06 (2006. 01)

H04L 1/16 (2006. 01)

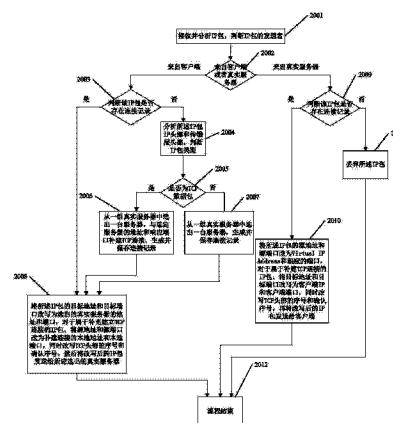
权利要求书3页 说明书5页 附图2页

## (54) 发明名称

一种在故障切换时保持 TCP 连接的方法

## (57) 摘要

本发明涉及一种在故障切换时保持 TCP 连接的方法,在主调度服务器与备调度服务器上同时执行以下步骤:接收并分析 IP 包,对于来自客户端的不存在连接记录的 IP 包做预处理,预处理包括:若是 TCP 数据包,从一组真实服务器中选出一台服务器,与选定服务器的地址和响应端口补建 TCP 连接,生成并保存连接记录,若不是 TCP 数据包,从一组真实服务器中选出一台服务器,生成并保存连接记录;然后将来自客户端 IP 包的目标地址和目标端口改写为选出的真实服务器的地址和端口,对于属于补建 TCP 连接的 IP 包,还要将源地址和源端口改为补建连接的本地地址和本地端口,同时改写 TCP 头部的序号和确认序号,再将改写后的 IP 包发送给所述选出的真实服务器。



1. 一种在故障切换时保持 TCP 连接的方法,在主调度服务器与备调度服务器上同时执行以下步骤:

接收并分析 IP 包,对于来自客户端的不存在连接记录的 IP 包做预处理,所述预处理包括:若是 TCP 数据包,从一组真实服务器中选出一台服务器,与选定服务器的地址和响应端口补建 TCP 连接,生成并保存连接记录,若不是 TCP 数据包,从一组真实服务器中选出一台服务器,生成并保存连接记录;然后将来自客户端 IP 包的目标地址和目标端口改写为选出的真实服务器的地址和端口,对于属于补建 TCP 连接的 IP 包,还要将源地址和源端口改为补建连接的本地地址和本地端口,同时改写 TCP 头部的序号和确认序号,再将改写后的 IP 包发送给所述选出的真实服务器。

2. 根据权利要求 1 所述的在故障切换时保持 TCP 连接的方法,其特征在于,还包括:对于来自真实服务器的 IP 包,若存在连接记录,将所述 IP 包的源地址和源端口改为虚拟 IP 地址和相应的端口,对于属于补建 TCP 连接的 IP 包,还要将目标地址和目标端口改写为客户端 IP 和客户端端口,同时改写 TCP 头部的序号和确认序号,再将改写后的 IP 包发送给客户端;若不存在连接记录,丢弃所述 IP 包。

3. 根据权利要求 2 所述的在故障切换时保持 TCP 连接的方法,其特征在于,该方法具体包括:

步骤 1)、主调度服务器或备调度服务器接收并分析 IP 包,判断 IP 包的发送者;

步骤 2)、若所述 IP 包来自客户端,执行步骤 3),若所述 IP 包来自真实服务器,执行步骤 9);

步骤 3)、检查连接记录表,判断所述 IP 包是否存在连接记录,若存在,执行步骤 8),否则执行步骤 4);其中,所述连接记录表中的每一条记录至少包括:客户端 IP、客户端端口、补建连接本地 IP、补建连接本地端口、真实服务器 IP、真实服务器端口、序号差、确认序号差;

步骤 4)、分析所述 IP 包的 IP 头部和传输层头部,判断 IP 包类型;

步骤 5)、若所述 IP 包为 TCP 数据包,执行步骤 6),否则,执行步骤 7);

步骤 6)、从一组真实服务器中选出一台服务器,然后将主调度服务器或备调度服务器作为客户端与本步骤中所选出服务器的地址和相应端口补充建立 TCP 连接,生成一条新的连接记录,并将新生成的连接记录保存到连接记录表中,接着执行步骤 8);

步骤 7)、从一组真实服务器中选出一台服务器,然后生成一条新的连接记录,并将新生成的连接记录保存到连接记录表中,接着执行步骤 8);

步骤 8)、将所述 IP 包的目标地址和目标端口改写为所述连接记录中所包含的选出的真实服务器的地址和端口;若所述 IP 包属于补建 TCP 连接的 IP 包,则还需要将源地址和源端口改为补建 TCP 连接的本地地址和本地端口,同时改写 TCP 头部的序号和确认序号;最后将上述改写后的 IP 包发送给所述选出的真实服务器,执行步骤 12);

步骤 9)、检查连接记录表,判断来自真实服务器的 IP 包是否存在连接记录,若存在,执行步骤 10),否则,执行步骤 11);

步骤 10)、将所述 IP 包的源地址和源端口改为虚拟 IP 地址和相应的端口,并进一步判断所述 IP 包是否属于补建 TCP 连接的 IP 包,若属于,将目标地址和目标端口改写为客户端 IP 和客户端端口,同时改写 TCP 头部的序号和确认序号,然后把改写后的 IP 包发给客户端,

再执行步骤 12), 否则, 直接把改写后的 IP 包发给客户端, 然后执行步骤 12);

步骤 11)、丢弃所述 IP 包;

步骤 12)、流程结束。

4. 根据权利要求 3 所述的在故障切换时保持 TCP 连接的方法, 其特征在于, 在所述的步骤 1) 中, 判断 IP 包的发送者包括: 若 IP 目标地址为虚拟 IP 地址, 则所述 IP 包来自客户端, 若源地址为真实服务器, 则所述 IP 包来自真实服务器。

5. 根据权利要求 3 所述的在故障切换时保持 TCP 连接的方法, 其特征在于, 在所述的步骤 3) 中, 判断所述 IP 包是否存在连接记录包括: 当所述来自客户端的 IP 包的源 IP、源端口与连接记录表中客户端 IP、客户端端口相等时, 表示所述 IP 包对应在连接记录表中存在记录。

6. 根据权利要求 3 所述的在故障切换时保持 TCP 连接的方法, 其特征在于, 在所述的步骤 6) 或步骤 7) 中, 从一组真实服务器中选出一台服务器的方法包括以下任意一种: 轮循算法、加权轮循算法、目的地址哈希算法、源地址哈希算法、最小连接算法。

7. 根据权利要求 3 所述的在故障切换时保持 TCP 连接的方法, 其特征在于, 在所述的步骤 6) 中, 生成一条新的连接记录包括: 将所述 IP 包的源 IP 和源端口作为连接记录中的客户端 IP 和客户端端口, 将本步骤中补建连接的本地 IP 和本地端口作为所述连接记录中的补建连接本地 IP 和补建连接本地端口, 将本步骤中所选择的真实服务器的地址和端口作为所述连接记录中的真实服务器 IP 和真实服务器端口, 将所述 IP 包的 TCP 头部的 SEQ 字段减去补建连接建立过程中三次握手最后一个确认包的 SEQ 作为所述连接记录中的序号差, 将所述 IP 包的 TCP 头部的 ACK 字段减去补建连接建立过程中三次握手最后一个确认包作为所述连接记录中的 ACK 确认序号差。

8. 根据权利要求 3 所述的在故障切换时保持 TCP 连接的方法, 其特征在于, 在所述的步骤 7) 中, 生成一条新的连接记录包括: 将所述 IP 包的源 IP 和源端口作为所述连接记录中的客户端 IP 和客户端端口, 将所述连接记录中的补建连接本地 IP 和补建连接本地端口设置为空, 将本步骤中所选择的真实服务器的地址和端口作为所述连接记录中的真实服务器 IP 和真实服务器端口, 将所述连接记录中的序号差设置为零, 将所述连接记录中的确认序号差设置为零。

9. 根据权利要求 3 所述的在故障切换时保持 TCP 连接的方法, 其特征在于, 在所述的步骤 8) 中, 改写 TCP 头部的序号包括: 将 TCP 头部的序号改为原 TCP 头部序号减连接记录中的序号差; 改写 TCP 头部的确认序号包括: 将 TCP 头部的确认序号改为原 TCP 头部确认序号减连接记录中的确认序号差。

10. 根据权利要求 3 所述的在故障切换时保持 TCP 连接的方法, 其特征在于, 在所述的步骤 10) 中, 改写 TCP 头部的序号包括: 将 TCP 头部的序号改为原 TCP 头部序号加上连接记录中的确认序号差; 改写 TCP 头部的确认序号包括: 将 TCP 头部的确认序号改为原 TCP 头部确认序号加上连接记录中的序号差。

11. 根据权利要求 3 所述的在故障切换时保持 TCP 连接的方法, 其特征在于, 在所述的步骤 9) 中, 判断来自真实服务器的 IP 包是否存在连接记录包括: 当所述 IP 包的源 IP、源端口、目的 IP、目的端口与连接记录表中真实服务器 IP、真实服务器端口、客户端 IP、客户端端口相等时, 表示所述 IP 包在连接记录表中存在记录; 当所述 IP 包的源 IP、源端口、目

的 IP、目的端口与连接记录表中真实服务器 IP、真实服务器端口、补建连接本地 IP、补建连接本地端口相等时,表示所述 IP 包在连接记录表中存在记录,并且该连接为补建 TCP 连接。

## 一种在故障切换时保持 TCP 连接的方法

### 技术领域

[0001] 本发明涉及通信技术领域,特别涉及在故障切换时保持 TCP 连接的方法。

### 背景技术

[0002] 故障切换(failover)指在发生故障时提供一种保护机制,当一台服务器失效时,另一台服务器会立刻接替工作,故障切换通过冗余来实现高可用性。

[0003] Linux Virtual Server (简称 LVS) 是基于 Linux 服务器集群解决方案,如图 1 所示,以 VS/NAT (Virtual Server via NAT) 模式为例,客户端 100 通过 Virtual IP Address (虚拟服务的 IP 地址) 访问网络服务时,请求报文到达调度服务器(假设为主调度服务器 102),调度服务器从一组真实服务器中选出一台服务器,例如真实服务器 105,将报文的目标地址 Virtual IP Address 改写成选定服务器的地址,报文的目标端口改写成选定服务器的相应端口,最后将修改后的报文发送给选出的服务器,同时调度服务器在连接 Hash 表中记录这个连接,当这个连接的下一个报文到达时,从连接 Hash 表中可以得到原选定服务器的地址和端口,进行同样的改写操作,并将报文传给原选定的服务器,当来自真实服务器的响应报文经过调度服务器时,调度服务器将报文的源地址和源端口改为 Virtual IP Address 和相应的端口,再把报文发给客户端。

[0004] 以 LVS 架设的主备调度服务器实现的高可用性为例,当主调度服务器 102 故障时,备调度服务器 103 通过心跳检测到主调度服务器故障,然后自动接管主调度服务器的工作,用户只感觉到一个虚拟服务器。

[0005] 然后现有技术中的上述故障切换方法存在以下缺点:

[0006] 在故障切换时,当与某一已建立连接属于同一 TCP 连接的报文被备调度服务器均衡到不同的真实服务器时,该真实服务器接收到报文后,并不知道数据报文段中提到的连接, TCP 处理的原则是接收方以复位作为应答,客户端将收到 RST 报文,造成连接被异常终止。因此对已经建立连接的 TCP 连接,在故障切换时,容易导致客户端 TCP 连接异常终止。

### 发明内容

[0007] 本发明的目的在于克服现有技术中的故障切换方法容易导致客户端 TCP 连接异常终止的缺陷,从而提供一种在故障切换时保持 TCP 连接的方法。

[0008] 为了实现上述目的,本发明提供了一种在故障切换时保持 TCP 连接的方法,在主调度服务器与备调度服务器上同时执行以下步骤:

[0009] 接收并分析 IP 包,对于来自客户端的不存在连接记录的 IP 包做预处理,所述预处理包括:若是 TCP 数据包,从一组真实服务器中选出一台服务器,与选定服务器的地址和响应端口补建 TCP 连接,生成并保存连接记录,若不是 TCP 数据包,从一组真实服务器中选出一台服务器,生成并保存连接记录;然后将来自客户端 IP 包的目标地址和目标端口改写为选出的真实服务器的地址和端口,对于属于补建 TCP 连接的 IP 包,还要将源地址和源端口改为补建连接的本地地址和本地端口,同时改写 TCP 头部的序号和确认序号,再将改写后

的 IP 包发送给所述选出的真实服务器。

[0010] 上述技术方案中,还包括:对于来自真实服务器的 IP 包,若存在连接记录,将所述 IP 包的源地址和源端口改为虚拟 IP 地址和相应的端口,对于属于补建 TCP 连接的 IP 包,还要将目标地址和目标端口改写为客户端 IP 和客户端端口,同时改写 TCP 头部的序号和确认序号,再将改写后的 IP 包发送给客户端;若不存在连接记录,丢弃所述 IP 包。

[0011] 上述技术方案中,该方法具体包括:

[0012] 步骤 1)、主调度服务器或备调度服务器接收并分析 IP 包,判断 IP 包的发送者;

[0013] 步骤 2)、若所述 IP 包来自客户端,执行步骤 3),若所述 IP 包来自真实服务器,执行步骤 9);

[0014] 步骤 3)、检查连接记录表,判断所述 IP 包是否存在连接记录,若存在,执行步骤 8),否则执行步骤 4);其中,所述连接记录表中的每一条记录至少包括:客户端 IP、客户端端口、补建连接本地 IP、补建连接本地端口、真实服务器 IP、真实服务器端口、序号差、确认序号差;

[0015] 步骤 4)、分析所述 IP 包的 IP 头部和传输层头部,判断 IP 包类型;

[0016] 步骤 5)、若所述 IP 包为 TCP 数据包,执行步骤 6),否则,执行步骤 7);

[0017] 步骤 6)、从一组真实服务器中选出一台服务器,然后将主调度服务器或备调度服务器作为客户端与本步骤中所选出服务器的地址和相应端口补充建立 TCP 连接,生成一条新的连接记录,并将新生成的连接记录保存到连接记录表中,接着执行步骤 8);

[0018] 步骤 7)、从一组真实服务器中选出一台服务器,然后生成一条新的连接记录,并将新生成的连接记录保存到连接记录表中,接着执行步骤 8);

[0019] 步骤 8)、将所述 IP 包的目标地址和目标端口改写为所述连接记录中所包含的选出的真实服务器的地址和端口;若所述 IP 包属于补建 TCP 连接的 IP 包,则还需要将源地址和源端口改为补建 TCP 连接的本地地址和本地端口,同时改写 TCP 头部的序号和确认序号;最后将上述改写后的 IP 包发送给所述选出的真实服务器,执行步骤 12);

[0020] 步骤 9)、检查连接记录表,判断来自真实服务器的 IP 包是否存在连接记录,若存在,执行步骤 10),否则,执行步骤 11);

[0021] 步骤 10)、将所述 IP 包的源地址和源端口改为虚拟 IP 地址和相应的端口,并进一步判断所述 IP 包是否属于补建 TCP 连接的 IP 包,若属于,将目标地址和目标端口改写为客户端 IP 和客户端端口,同时改写 TCP 头部的序号和确认序号,然后把改写后的 IP 包发给客户端,再执行步骤 12),否则,直接把改写后的 IP 包发给客户端,然后执行步骤 12);

[0022] 步骤 11)、丢弃所述 IP 包;

[0023] 步骤 12)、流程结束。

[0024] 上述技术方案中,在所述的步骤 1)中,判断 IP 包的发送者包括:若 IP 目标地址为虚拟 IP 地址,则所述 IP 包来自客户端,若源地址为真实服务器,则所述 IP 包来自真实服务器。

[0025] 上述技术方案中,在所述的步骤 3)中,判断所述 IP 包是否存在连接记录包括:当所述来自客户端的 IP 包的源 IP、源端口与连接记录表中客户端 IP、客户端端口相等时,表示所述 IP 包对应存在连接记录表中存在记录。

[0026] 上述技术方案中,在所述的步骤 6)或步骤 7)中,从一组真实服务器中选出一台服

服务器的方法包括以下任意一种：轮循算法、加权轮循算法、目的地址哈希算法、源地址哈希算法、最小连接算法。

[0027] 上述技术方案中，在所述的步骤 6) 中，生成一条新的连接记录包括：将所述 IP 包的源 IP 和源端口作为连接记录中的客户端 IP 和客户端端口，将本步骤中补建连接的本地 IP 和本地端口作为所述连接记录中的补建连接本地 IP 和补建连接本地端口，将本步骤中所选择的真实服务器的地址和端口作为所述连接记录中的真实服务器 IP 和真实服务器端口，将所述 IP 包的 TCP 头部的 SEQ 字段减去补建连接建立过程中三次握手最后一个确认包的 SEQ 作为所述连接记录中的序号差，将所述 IP 包的 TCP 头部的 ACK 字段减去补建连接建立过程中三次握手最后一个确认包作为所述连接记录中的 ACK 确认序号差。

[0028] 上述技术方案中，在所述的步骤 7) 中，生成一条新的连接记录包括：将所述 IP 包的源 IP 和源端口作为所述连接记录中的客户端 IP 和客户端端口，将所述连接记录中的补建连接本地 IP 和补建连接本地端口设置为空，将本步骤中所选择的真实服务器的地址和端口作为所述连接记录中的真实服务器 IP 和真实服务器端口，将所述连接记录中的序号差设置为零，将所述连接记录中的确认序号差设置为零。

[0029] 上述技术方案中，在所述的步骤 8) 中，改写 TCP 头部的序号包括：将 TCP 头部的序号改为原 TCP 头部序号减连接记录中的序号差；改写 TCP 头部的确认序号包括：将 TCP 头部的确认序号改为原 TCP 头部确认序号减连接记录中的确认序号差。

[0030] 上述技术方案中，在所述的步骤 10) 中，改写 TCP 头部的序号包括：将 TCP 头部的序号改为原 TCP 头部序号加上连接记录中的确认序号差；改写 TCP 头部的确认序号包括：将 TCP 头部的确认序号改为原 TCP 头部确认序号加上连接记录中的序号差。

[0031] 上述技术方案中，在所述的步骤 9) 中，判断来自真实服务器的 IP 包是否存在连接记录包括：当所述 IP 包的源 IP、源端口、目的 IP、目的端口与连接记录表中真实服务器 IP、真实服务器端口、客户端 IP、客户端端口相等时，表示所述 IP 包在连接记录表中存在记录；当所述 IP 包的源 IP、源端口、目的 IP、目的端口与连接记录表中真实服务器 IP、真实服务器端口、补建连接本地 IP、补建连接本地端口相等时，表示所述 IP 包在连接记录表中存在记录，并且该连接为补建 TCP 连接。

[0032] 本发明的优点在于：

[0033] 本发明的方法在故障切换时可以保持已经建立连接成功的 TCP 连接，避免客户端 TCP 连接异常终止，并且在整个故障切换过程中，对客户端是透明的。

## 附图说明

[0034] 图 1 是 Linux 服务器集群的示意图；

[0035] 图 2 是本发明的在故障切换时保持 TCP 连接的方法的流程图；

[0036] 图 3 是连接记录表的示意图。

## 具体实施方式

[0037] 现结合附图对本发明作进一步的描述。

[0038] 在一个实施例中，仍以图 1 中所示的 Linux 服务器集群为例，对本发明的在故障切换时保持 TCP 连接的方法进行描述。

[0039] 从背景技术部分的描述可以知道：当主调度服务器 102 失效后，备调度服务器 103 通过 HeartBeat 心跳监测到主调度服务器 102 失效，然后接管主调度服务器 102 的任务，当主调度服务器 102 恢复后再重新接管备调度服务器 103 的任务。本发明的方法需要在主调度服务器 102 和备调度服务器 103 上同时实施，从而在备调度服务器 103 接管主调度服务器 102 任务以及主调度服务器 102 恢复后接管备调度服务器 103 任务时，保持已经建立的 TCP 连接。

[0040] 参考图 2，下面对本发明的方法做详细说明。

[0041] 步骤 2001)、主(或备)调度服务器利用 netfilter 接收 IP 包，分析 IP 包的 IP 头部参数，根据分析结果判断 IP 包的发送者：若 IP 目标地址为 Virtual IP Address，则所述 IP 包来自客户端，若源地址为真实服务器，则所述 IP 包来自真实服务器；

[0042] 步骤 2002)、根据 IP 包的来源，执行不同的操作：若 IP 包来自客户端，转步骤 2003)，若 IP 包来自真实服务器，转步骤 2009)；

[0043] 步骤 2003)、检查连接记录表，判断该 IP 包是否存在连接记录，若存在转步骤 2008)，否则转步骤 2004)；

[0044] 在本步骤中，所述连接记录表如图 3 所述，每条记录包括且不限于：客户端 IP、客户端端口、补建连接本地 IP、补建连接本地端口、真实服务器 IP、真实服务器端口、序号差、确认序号差。

[0045] 本步骤中，当所述来自客户端的 IP 包的源 IP、源端口与连接记录表中客户端 IP、客户端端口相等时，表示所述 IP 包对应存在连接记录表中存在记录。

[0046] 步骤 2004)、分析所述 IP 包的 IP 头部和传输层头部，判断 IP 包类型：若传输层为 TCP 且 TCP 载荷不为空，则为 TCP 数据包；

[0047] 步骤 2005)、根据 IP 包的类型执行不同的操作：若为 TCP 数据包，转步骤 2006)，否则，转步骤 2007)；

[0048] 步骤 2006)、主(或备)调度服务器从一组真实服务器中选出一台服务器，然后主(或备)调度服务器作为客户端与选定服务器的地址和相应端口补充建立 TCP 连接，生成一条新的连接记录，并将新生成的连接记录保存到连接记录表中，接着执行步骤 2008)。

[0049] 在本步骤中，所述主(或备)调度服务器从一组真实服务器中选出一台服务器的选择算法可以有多种，包括但不限于轮循算法、加权轮循算法、目的地址哈希算法、源地址哈希算法、最小连接算法。

[0050] 在本步骤中，连接记录包括但不限于：客户端 IP、客户端端口、补建连接本地 IP、补建连接本地端口、真实服务器 IP、真实服务器端口、序号差、确认序号。具体的说，其中的客户端 IP 和客户端端口为之前所述 IP 包的源 IP 和源端口，补建连接本地 IP 和补建连接本地端口为本步骤中补建连接的本地 IP 和本地端口，真实服务器 IP 和真实服务器端口为选择的真实服务器的地址和端口，序号差为所述 IP 包的 TCP 头部的 SEQ 字段减去补建连接建立过程中三次握手最后一个确认包的 SEQ，确认序号差为所述 IP 包的 TCP 头部的 ACK 字段减去补建连接建立过程中三次握手最后一个确认包的 ACK。

[0051] 步骤 2007)、主(或备)调度服务器从一组真实服务器中选出一台服务器，然后生成一条新的连接记录，并将新生成的连接记录保存到连接记录表中。

[0052] 在本步骤中，所述主(或备)调度服务器从一组真实服务器中选出一台服务器的选



择算法可以有多种,包括但不限于轮循算法、加权轮循算法、目的地址哈希算法、源地址哈希算法、最小连接算法等。

[0053] 所述新生成的连接记录中的参数值如下:客户端 IP 和客户端端口为所述 IP 包的源 IP 和源端口,补建连接本地 IP 和补建连接本地端口为空,真实服务器 IP 和真实服务器端口为选择的真实服务器的地址和端口,序号差为零,确认序号差为零。

[0054] 步骤 2008)、将所述 IP 包的目标地址和目标端口改写为所述连接记录中所包含的选出的真实服务器的地址和端口;若所述 IP 包属于补建 TCP 连接的 IP 包,则还需要将源地址和源端口改为补建 TCP 连接的本地地址和本地端口,同时改写 TCP 头部的序号和确认序号,最后将上述改写后的 IP 包发送给所述选出的真实服务器,转步骤 2012);

[0055] 本步骤中,改写 TCP 头部的序号包括:将 TCP 头部的序号改为原 TCP 头部序号减连接记录中的序号差;改写 TCP 头部的确认序号包括:将 TCP 头部的确认序号改为原 TCP 头部确认序号减连接记录中的确认序号差。

[0056] 步骤 2009)、检查连接记录表,判断来自真实服务器的 IP 包是否存在连接记录,若存在,执行步骤 2010),否则,执行步骤 2011);

[0057] 本步骤中,当所述 IP 包的源 IP、源端口、目的 IP、目的端口与连接记录表中真实服务器 IP、真实服务器端口、客户端 IP、客户端端口相等时,表示所述 IP 包在连接记录表中存在记录;当所述 IP 包的源 IP、源端口、目的 IP、目的端口与连接记录表中真实服务器 IP、真实服务器端口、补建连接本地 IP、补建连接本地端口相等时,表示所述 IP 包在连接记录表中存在记录,并且该连接为补建 TCP 连接。

[0058] 步骤 2010)、将所述 IP 包的源地址和源端口改为 Virtual IP Address(虚拟 IP 地址)和相应的端口,并进一步判断所述 IP 包是否属于补建 TCP 连接的 IP 包,若属于,将目标地址和目标端口改写为客户端 IP 和客户端端口,同时改写 TCP 头部的序号和确认序号,然后把改写后的 IP 包发给客户端,再执行步骤 2012),否则,直接把改写后的 IP 包发给客户端,然后执行步骤 2012);

[0059] 本步骤中,改写 TCP 头部的序号包括:将 TCP 头部的序号改为原 TCP 头部序号加上连接记录中的确认序号差;将 TCP 头部的确认序号改为原 TCP 头部确认序号加上连接记录中的序号差。

[0060] 步骤 2011)、丢弃所述 IP 包;

[0061] 步骤 2012)、流程结束。

[0062] 上述描述以 Linux 服务器集群为例,对本发明的方法做了说明。在其他实施例中,本发明的方法也适用于其他类型的服务器集群。

[0063] 最后所应说明的是,以上实施例仅用以说明本发明的技术方案而非限制。尽管参照实施例对本发明进行了详细说明,本领域的普通技术人员应当理解,对本发明的技术方案进行修改或者等同替换,都不脱离本发明技术方案的精神和范围,其均应涵盖在本发明的权利要求范围当中。

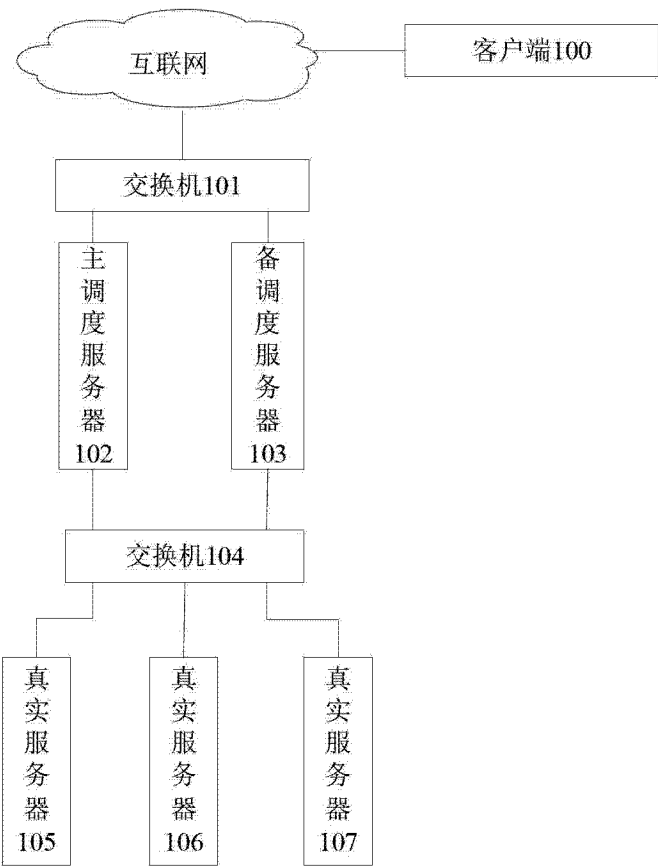


图 1

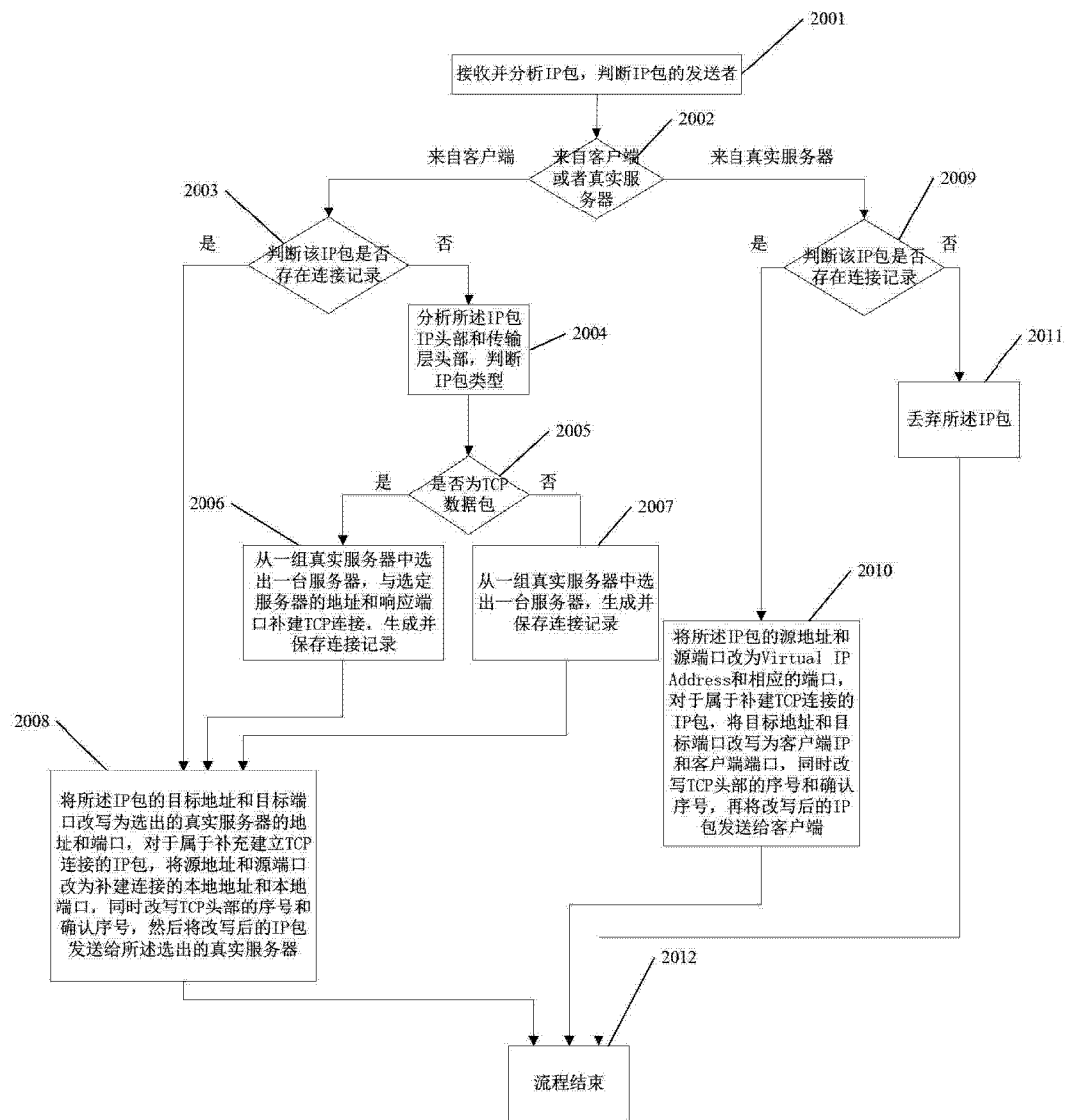


图 2

客户端IP	客户端端口	补建连接本地IP	补建连接本地端口	真实服务器IP	真实服务器端口	序号差	确认序号差
-------	-------	----------	----------	---------	---------	-----	-------

图 3