



(12)发明专利申请

(10)申请公布号 CN 108322402 A

(43)申请公布日 2018.07.24

(21)申请号 201710039457.7

(22)申请日 2017.01.18

(71)申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 耿雪松 杜宗鹏 董杰

(74)专利代理机构 深圳市深佳知识产权代理事务所(普通合伙) 44285

代理人 王仲凯

(51) Int. Cl.

H04L 12/801(2013.01)

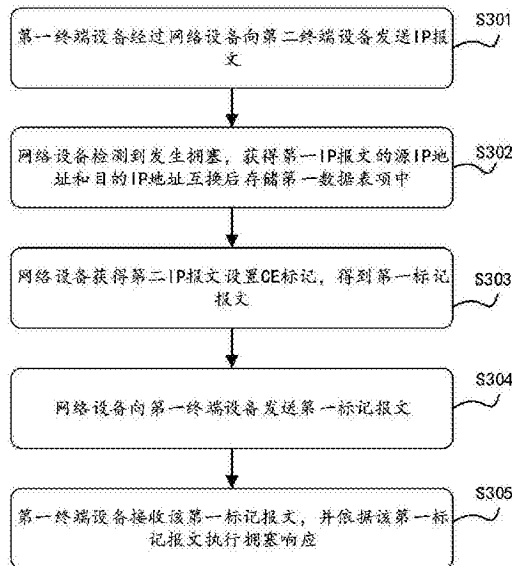
权利要求书4页 说明书18页 附图5页

(54)发明名称

报文处理方法、设备及系统

(57)摘要

本申请提供了一种报文处理方法、设备及系统,当网络设备发生拥塞时,获得来自第一终端设备的第一IP报文的源IP地址和目的IP地址,进行互换后存储于第一数据表项中的源IP地址和目的IP地址中,再获得来自第二终端设备的第二IP报文,该第二IP报文的源IP地址与第一数据表项中的源IP地址相同,目的IP地址与第一数据表项中的目的IP地址相同,对第二IP报文设置CE标记,得到第一标记报文,并向第一终端设备发送该第一标记报文。在拥塞探测时,不需要发送给对端设备进行响应,也不需要依赖终端设备自身判断对应该报文的计时是否超时,从而实现提高拥塞探测的速率,降低延时的目的。



1. 一种报文处理方法,其特征在于,所述方法包括:

当网络设备发生拥塞时,所述网络设备获得来自第一终端设备的第一因特网协议IP报文的源IP地址和目的IP地址,将所述第一IP报文的源IP地址存储于第一数据表项的目的IP地址中,将所述第一IP报文的源IP地址存储于所述第一数据表项的源IP地址中;

所述网络设备获得来自第二终端设备的第二IP报文,并确定所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的目的IP地址相同,对所述第二IP报文设置拥塞发生CE标记,得到第一标记报文,所述CE标记用于标识所述网络设备在传输报文时发生了拥塞;

所述网络设备向所述第一终端设备发送所述第一标记报文。

2. 根据权利要求1所述的方法,其特征在于,所述第一数据表项中还包括计数器,所述计数器用于指示待设置CE标记的IP报文的个数,所述待设置CE标记的IP报文的源IP地址和目的IP地址与所述第二IP报文的源IP地址和目的IP地址分别相同;

相应地,获得所述待设置CE标记的IP报文的个数,包括:

所述网络设备每接收到一个与所述第一IP报文的源IP地址和目的IP地址分别相同的IP报文时,将所述计数器加1;

所述网络设备每接收到一个与所述第二IP报文的源IP地址和目的IP地址分别相同的IP报文时,将所述计数器减1。

3. 根据权利要求2所述的报文处理方法,其特征在于,所述第一数据表项中还包括计时器;

所述网络设备在生成所述第一数据表项或对所述第一数据表项中的计数器减1时,所述计时器重置计时;

若所述计时器的计时达到预设计时,所述网络设备删除所述第一数据表项。

4. 根据权利要求1所述的报文处理方法,其特征在于,所述第一数据表项中还包括源端口号和目的端口号;

所述网络设备获得来自所述第一终端设备的第一IP报文中的源端口号和目的端口号,将所述第一IP报文的源端口号存储于所述第一数据表项的目的端口号中,将所述第一IP报文的源端口号存储于所述第一数据表项的源端口号中;

相应地,所述网络设备确定所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的目的IP地址相同,对所述第二IP报文设置拥塞发生CE标记,包括:

确定所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的目的IP地址相同,所述第二IP报文的源端口号与所述第一数据表项中的源端口号相同,所述第二IP报文的源端口号与所述第一数据表项中的目的端口号相同,对所述第二IP报文设置拥塞发生CE标记。

5. 根据权利要求2-4中任意一项所述的报文处理方法,其特征在于,所述待设置CE标记的IP报文的源IP地址、目的IP地址、源端口号和目的端口号与所述第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同;

相应地,获得所述待设置CE标记的IP报文的个数,包括:

所述网络设备每接收到一个与所述第一IP报文的源IP地址、目的IP地址、源端口号和

目的端口号分别相同的IP报文时,将所述计数器加1;

所述网络设备每接收到一个与所述第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同的IP报文时,将所述计数器减1。

6. 根据权利要求1-5中任意一项所述的报文处理方法,其特征在于,在所述网络设备获得第一IP报文之后,所述方法还包括:

所述网络设备对所述第一IP报文添加标识信息,所述标识信息用于指示所述网络设备根据所述第一IP报文携带的信息生成所述第一数据表项和/或对所述第一数据表项中的计数器加1;所述第一IP报文携带的信息包括所述第一IP报文的源IP地址和目的IP地址,或者所述第一IP报文的源IP地址、目的IP地址、源端口号和目的端口号。

7. 根据权利要求1-6中任意一项所述的报文处理方法,其特征在于,所述第一IP报文的IP报文头中包括反向显式拥塞通告RECN标记,所述RECN标记用于标识所述第一IP报文所属的数据流支持反向拥塞探测。

8. 一种报文处理方法,其特征在于,适用于第一终端设备,若所述第一终端设备发送的数据流支持反向拥塞探测机制,所述方法包括:

第一终端设备向网络设备发送第一因特网协议IP报文,所述第一IP报文的IP报文头中包括反向显式拥塞通告RECN标记,所述RECN标记用于标识所述第一IP报文所属的数据流支持反向拥塞探测机制;

所述第一终端设备获得来自所述网络设备的第一标记报文,所述第一标记报文包括拥塞发生CE标记,所述CE标记用于标识所述网络设备在传输报文时发生了拥塞;

所述第一终端设备基于所述第一标记报文执行拥塞响应。

9. 根据权利要求8所述的报文处理方法,其特征在于,若所述第一终端设备发送的数据流支持反向拥塞探测机制和正向拥塞探测机制,还包括:

所述第一终端设备获得来自所述网络设备的第二标记报文,所述第二标记报文包括所述CE标记;

所述第一终端设备确定所述第二标记报文的源端口号和目的端口号,基于所述第二标记报文的源端口号和目的端口号确定所述第二标记报文所属的数据流;

若所述第一终端设备确定所述数据流所支持的拥塞探测机制为反向拥塞探测机制,所述第一终端设备执行对所述数据流的拥塞处理;

若所述第一终端设备确定所述数据流所支持的拥塞探测为正向拥塞探测机制,则向所述网络设备发送响应ACK消息,所述ACK消息用于指示第二终端设备执行对所述数据流的拥塞处理,所述ACK消息中包括CE标记。

10. 一种网络设备,其特征在于,包括:

处理单元,用于若所述网络设备发生拥塞时,获得来自第一终端设备的第一因特网协议IP报文的源IP地址和目的IP地址,将所述第一IP报文的源IP地址存储于第一数据表项的目的IP地址中,将所述第一IP报文的源IP地址存储于所述第一数据表项的源IP地址中,以及获得来自第二终端设备的第二IP报文,所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的目的地址相同,对所述第二IP报文设置拥塞发生CE标记,得到第一标记报文,所述CE标记用于标识所述网络设备在传输报文时发生了拥塞;

发送单元,用于向所述第一终端设备发送所述第一标记报文。

11. 根据权利要求10所述的网络设备,其特征在于,所述第一数据表项中还包括计数器,所述计数器用于指示待设置CE标记的IP报文的个数,所述待设置CE标记的IP报文的源IP地址和目的IP地址与所述第二IP报文的源IP地址和目的IP地址分别相同;

相应地,所述处理单元,用于每接收到一个与所述第一IP报文的源IP地址和目的IP地址分别相同的IP报文时,将所述计数器加1,以及每得到一个与所述第二报文的源IP地址和目的IP地址分别相同的IP报文时,将所述计数器减1。

12. 根据权利要求11所述的网络设备,其特征在于,所述第一数据表项中还包括计时器;

所述处理单元,还用于在生成所述第一数据表项或对所述第一数据表项中的计数器减1时,所述计时器重置计时,若所述计时器的计时达到预设计时,删除所述第一数据表项。

13. 根据权利要求10所述的网络设备,其特征在于,所述第一数据表项中还包括源端口号和目的端口号;

所述处理单元,还用于获取来自所述第一终端设备的第一IP报文中的源端口号和目的端口号,将所述第一IP报文的源端口号存储于所述第一数据表项的目的端口号中,将所述第一IP报文的源端口号存储于所述第一数据表项的源端口号中;以及确定所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的目的地址相同,确定所述第二IP报文的源端口号与所述第一数据表项中的源端口号相同,所述第二IP报文的源端口号与所述第一数据表项中的目的端口号相同,对所述第二IP报文设置拥塞发生CE标记。

14. 根据权利要求11-13中任意一项所述的网络设备,其特征在于,所述待设置CE标记的IP报文的源IP地址、目的IP地址、源端口号和目的端口号与所述第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同;

相应地,所述处理单元,用于每接收到一个与所述第一IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同的IP报文时,将所述计数器加1;每接收到一个与所述第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同的IP报文时,将所述计数器减1。

15. 根据权利要求10-14中任意一项所述的网络设备,其特征在于,

所述处理单元,还用于在获得所述第一IP报文之后,对所述第一IP报文添加标识信息,所述标识信息用于指示所述网络设备根据所述第一IP报文携带的信息生成所述第一数据表项和/或对所述第一数据表项中的计数器加1;所述第一IP报文携带的信息包括所述第一IP报文的源IP地址和目的IP地址,或者所述第一IP报文的源IP地址、目的IP地址,源端口号和目的端口号。

16. 一种终端设备,用作第一终端设备,其特征在于,若所述第一终端设备发送的数据流支持反向拥塞探测机制,所述第一终端设备包括:

发送单元,用于向网络设备发送第一因特网协议IP报文,所述第一IP报文的IP报文头中包括反向显式拥塞通告RECN标记,所述RECN标记用于标识所述第一IP报文所属的数据流支持反向拥塞探测;

接收单元,用于接收来自所述网络设备的第一标记报文,所述第一标记报文包括拥塞

发生CE标记,所述CE标记用于标识所述网络设备在传输报文时发生了拥塞;

处理单元,用于基于所述第一标记报文执行拥塞响应。

17. 根据权利要求16所述的终端设备,其特征在于,若所述第一终端设备发送的数据流支持反向拥塞探测机制和正向拥塞探测机制;

所述接收单元,还用于接收来自所述网络设备的第二标记报文,所述第二标记报文包括CE标记;

所述处理单元,还用于确定所述第二标记报文的源端口号和目的端口号,并基于所述第二标记报文的源端口号和目的端口号确定所述第二标记报文所属的数据流;若确定所述数据流所支持的拥塞探测机制为反向拥塞探测机制,则所述第一终端设备执行对所述数据流的拥塞处理;若确定所述数据流所支持的拥塞探测机制为正向拥塞探测机制,则向所述网络设备发送响应ACK消息,所述ACK消息用于指示第二终端设备执行对所述数据流的拥塞处理,所述ACK消息中包括CE标记。

18. 一种网络系统,其特征在于,包括:第一终端设备和网络设备;

所述第一终端设备为权利要求16或17所述的终端设备;

所述网络设备为权利要求10-15中任意一项所述的网络设备。

报文处理方法、设备及系统

技术领域

[0001] 本发明涉及通信技术领域,尤其是,涉及一种报文处理方法、设备及系统。

背景技术

[0002] 网络资源包括带宽、缓存空间和终端的处理能力等。在网络中依赖网络资源进行数据传输。一旦数据传输所需要的网络资源超出了当前可用的网络资源,网络会发生拥塞。因此,需要在数据传输的过程中对网络资源进行探测,并在数据流传输结束之前对探测到的网络拥塞进行响应,从而完成对网络拥塞的控制。

[0003] 在进行网络拥塞控制的过程中,常用的拥塞探测方法为:隐式拥塞通告方法。该隐式拥塞通告方法基于传输控制协议(英文:Transmission Control Protocol,TCP)的可靠连接,在发送端设备通过网络设备向接收端设备发送报文的过程中,每发送一个报文开启一个计时器。当发送端设备接收到接收端设备反馈的响应(英文:Acknowledgement,ACK)消息后,关闭计时器。当报文在网络设备处发生拥塞时,网络设备直接丢弃该报文。因接收端设备未接收到该报文,则不向发送端设备反馈ACK消息。发送端设备未接收到该ACK消息,对应报文的计时器超时,即判断报文丢失,网络发生拥塞,发送端设备进行拥塞响应。上述拥塞探测的时间取决于TCP协议中超时重传的时间,一般大于一个往返时间(英文:Round Trip Time,RTT),探测时间较长。

[0004] 目前数据网络中出现了越来越多短寿命数据流(英文:short-lived flow),且数据网络中承载的对传输时延敏感的业务也越来越多,因此需要更加快速灵敏的拥塞探测方法,以提升拥塞控制效率,降低时延。而现有技术中的拥塞探测方法耗时长,已经无法满足目前数据网络对于拥塞探测的要求。

发明内容

[0005] 有鉴于此,本申请实施例提供一种报文处理方法、设备及系统,用于解决现有技术中拥塞探测耗时长的问题。

[0006] 本申请实施例提供如下技术方案:

[0007] 本申请实施例第一方面公开了一种报文处理方法,所述方法包括:

[0008] 当网络设备发生拥塞时,所述网络设备获得来自第一终端设备的第一因特网协议IP报文的源IP地址和目的IP地址,将所述第一IP报文的源IP地址存储于第一数据表项的目的IP地址中,将所述第一IP报文的源IP地址存储于所述第一数据表项的源IP地址中;

[0009] 所述网络设备获得来自第二终端设备的第二IP报文,并确定所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的目的IP地址与所述第一数据表项中的目的IP地址相同,对所述第二IP报文设置拥塞发生CE标记,得到第一标记报文,所述CE标记用于标识所述网络设备在传输报文时发生了拥塞;

[0010] 所述网络设备向所述第一终端设备发送所述第一标记报文。

[0011] 上述方案,网络设备在检测到发生拥塞后,将获得来自第一终端设备的第一IP报

文的源IP地址和目的IP地址交换后,存储于第一数据表项中的源IP地址和目的IP地址中,基于此,获得与第一数据表项中的源IP地址和目的IP地址相同的第二IP报文,并设置CE标记,该第二IP报文为第二终端设备发送的。通过在反向传输的IP报文上设置CE标记,在拥塞探测时,不需要发送给对端设备进行响应,也不需要依赖终端设备自身判断对应该报文的计时是否超时,减少了拥塞探测的时间,进而降低时延。

[0012] 在一种可能的设计中,所述第一数据表项中还包括计数器,所述计数器用于指示待设置CE标记的IP报文的个数,所述待设置CE标记的IP报文的源IP地址和目的IP地址与所述第二IP报文的源IP地址和目的IP地址分别相同。

[0013] 获得所述待设置CE标记的IP报文的个数,包括:

[0014] 所述网络设备每接收到一个与所述第一IP报文的源IP地址和目的IP地址分别相同的IP报文时,将所述计数器加1;

[0015] 所述网络设备每接收到一个与所述第二IP报文的源IP地址和目的IP地址分别相同的IP报文时,将所述计数器减1。

[0016] 上述方案,网络设备在第一数据表项中增加计数器,可以确定待设置CE标记的IP报文的个数,若个数越多则指示第一终端设备需要采取的拥塞响应的效率越高。此外,网络设备通过对计数器中的计数进行动态控制,可以有效的得到待设置CE标记的IP报文的个数。

[0017] 在一种可能的设计中,所述第一数据表项中还包括计时器;

[0018] 所述网络设备在生成所述第一数据表项或对所述第一数据表项中的计数器减1时,所述计时器重置计时;

[0019] 若所述计时器的计时达到预设计时,所述网络设备删除所述第一数据表项。

[0020] 上述方案,网络设备在第一数据表项中增加计时器,为老化第一数据表项的过程,可以在未获得待设置CE标记的IP报文时,及时删除该第一数据表项。进一步的,还可以解决在负载均衡的情况下,正向数据流和反向数据流可以路径不同的问题。

[0021] 在一种可能的设计中,所述第一数据表项中还包括源端口号和目的端口号;

[0022] 所述网络设备获得来自所述第一终端设备的第一IP报文中的源端口号和目的端口号,将所述第一IP报文的源端口号存储于所述第一数据表项的目的端口号中,将所述第一IP报文的源端口号存储于所述第一数据表项的源端口号中;

[0023] 相应地,所述网络设备确定所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,对所述第二IP报文设置拥塞发生CE标记,包括:

[0024] 确定所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,对所述第二IP报文设置拥塞发生CE标记。

[0025] 上述方案,网络设备在第一数据表项中增加源端口号和目的端口号,可以解决相同的发送端设备和接收端设备之间存在不同数据流在网络设备处出现拥塞的情况。具有相同源端口号和目的端口号的数据流为同一数据流。

[0026] 在一种可能的设计中,所述待设置CE标记的IP报文的源IP地址、目的IP地址、源端

口号和目的端口号与所述第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同；

[0027] 相应地,获得所述待设置CE标记的IP报文的个数,包括:

[0028] 所述网络设备每接收到一个与所述第一IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同的IP报文时,将所述计数器加1;

[0029] 所述网络设备每接收到一个与所述第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同的IP报文时,将所述计数器减1。

[0030] 上述方案,网络设备在第一数据表项中增加计数器,可以确定待设置CE标记的IP报文的个数,若个数越多则指示第一终端设备需要采取的拥塞响应的效率越高。此外,网络设备通过对计数器中的计数进行动态控制,可以有效的得到待设置CE标记的IP报文的个数。

[0031] 在一种可能的设计中,在所述网络设备获得第一IP报文之后,还包括:

[0032] 所述网络设备对所述第一IP报文添加标识信息,所述标识信息用于指示所述网络设备根据所述第一IP报文携带的信息生成所述第一数据表项和/或对所述第一数据表项中的计数器加1;所述第一IP报文携带的信息包括所述第一IP报文的源IP地址和目的IP地址,或者所述第一IP报文的源IP地址、目的IP地址、源端口号和目的端口号。

[0033] 上述方案,网络设备对IP报文添加标识信息,有利于在当网络设备中存在上下行处理流程时,能够快速识别对哪些IP报文进行处理。

[0034] 在一种可能的设计中,所述第一IP报文的IP报文头中包括反向显式拥塞通告RECN标记,所述RECN标记用于标识所述第一IP报文所属的数据流支持反向拥塞探测。

[0035] 本申请实施例的第二方面公开了一种报文处理方法,适用于第一终端设备,若所述第一终端设备发送的数据流支持反向拥塞探测机制,所述方法包括:

[0036] 第一终端设备向网络设备发送第一因特网协议IP报文,所述第一IP报文的IP报文头中包括反向显式拥塞通告RECN标记,所述RECN标记用于标识所述第一IP报文所属的数据流支持反向拥塞探测机制;

[0037] 所述第一终端设备获得来自所述网络设备的第一标记报文,所述第一标记报文包括拥塞发生CE标记,所述CE标记用于标识所述网络设备在传输报文时发生了拥塞;

[0038] 所述第一终端设备基于所述第一标记报文执行拥塞响应。

[0039] 上述方案,第一终端设备在向网络设备发送的IP报文中添加RECN标记,使网络设备在接收到终端设备发送的IP报文,并基于其支持的RECN机制,通过在反向传输的IP报文中设置CE标记,然后反馈该终端设备,可以使终端设备直接通过该CE标记判断自身的报文传输发生拥塞。从而减少了拥塞探测的时间,进而降低时延。

[0040] 在一种可能的设计中,若所述第一终端设备发送的数据流支持反向拥塞探测机制和正向拥塞探测机制,还包括:

[0041] 所述第一终端设备获得来自所述网络设备的第二标记报文,所述第二标记报文包括所述CE标记;

[0042] 所述第一终端设备确定所述第二标记报文的源端口号和目的端口号,基于所述第二标记报文的源端口号和目的端口号确定所述第二标记报文所属的数据流;

[0043] 若所述第一终端设备确定所述数据流所支持的拥塞探测机制为反向拥塞探测机

制,所述第一终端设备执行对所述数据流的拥塞处理;

[0044] 若所述第一终端设备确定所述数据流所支持的拥塞探测为正向拥塞探测机制,则向所述网络设备发送响应ACK消息,所述ACK消息用于指示第二终端设备执行对所述数据流的拥塞处理,所述ACK消息中包括CE标记。

[0045] 上述方案,第一终端设备在发送报文时,基于在报文头中ECN FIELD的不同置位,区分各个数据流所支持的拥塞探测机制。网络设备根据接收到的报文中的报文头的ECN FIELD不同置位,执行相应地拥塞探测机制。基于不同数据流支持不同拥塞探测机制的方式,使网络设备在执行拥塞探测时,能够更加灵活。

[0046] 本申请实施例的第三方面公开了一种网络设备,包括:

[0047] 处理单元,用于若所述网络设备发生拥塞时,获得来自第一终端设备的第一因特网协议IP报文的源IP地址和目的IP地址,将所述第一IP报文的源IP地址存储于第一数据表项的目的IP地址中,将所述第一IP报文的源IP地址存储于所述第一数据表项的源IP地址中,以及获得来自第二终端设备的第二IP报文,所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的目的地址相同,对所述第二IP报文设置拥塞发生CE标记,得到第一标记报文,所述CE标记用于标识所述网络设备在传输报文时发生了拥塞;

[0048] 发送单元,用于向所述第一终端设备发送所述第一标记报文。

[0049] 在一种可能的设计中,所述第一数据表项中还包括计数器,所述计数器用于指示待设置CE标记的IP报文的个数,所述待设置CE标记的IP报文的源IP地址和目的IP地址与所述第二IP报文的源IP地址和目的IP地址分别相同;

[0050] 相应地,所述处理单元,用于每接收到一个与所述第一IP报文的源IP地址和目的IP地址分别相同的IP报文时,将所述计数器加1,以及每得到一个与所述第二报文的源IP地址和目的IP地址分别相同的IP报文时,将所述计数器减1。

[0051] 在一种可能的设计中,所述第一数据表项中还包括计时器;

[0052] 所述处理单元,还用于在生成所述第一数据表项或对所述第一数据表项中的计数器减1时,所述计时器重置计时,若所述计时器的计时达到预设计时,删除所述第一数据表项。

[0053] 在一种可能的设计中,所述第一数据表项中还包括源端口号和目的端口号;

[0054] 所述处理单元,还用于获取来自所述第一终端设备的第一IP报文中的源端口号和目的端口号,将所述第一IP报文的源端口号存储于所述第一数据表项的目的端口号中,将所述第一IP报文的目的端口号存储于所述第一数据表项的源端口号中;以及确定所述第二IP报文的源IP地址与所述第一数据表项中的源IP地址相同,所述第二IP报文的源IP地址与所述第一数据表项中的目的地址相同,确定所述第二IP报文的源端口号与所述第一数据表项中的源端口号相同,所述第二IP报文的源端口号与所述第一数据表项中的目的端口号相同,对所述第二IP报文设置拥塞发生CE标记。

[0055] 在一种可能的设计中,所述待设置CE标记的IP报文的源IP地址、目的IP地址、源端口号和目的端口号与所述第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同;

[0056] 相应地,所述处理单元,用于每接收到一个与所述第一IP报文的源IP地址、目的IP

地址、源端口号和目的端口号分别相同的IP报文时,将所述计数器加1;每接收到一个与所述第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同的IP报文时,将所述计数器减1。

[0057] 在一种可能的设计中,所述处理单元,还用于在获得所述第一IP报文之后,对所述第一IP报文添加标识信息,所述标识信息用于指示所述网络设备根据所述第一IP报文携带的信息生成所述第一数据表项和/或对所述第一数据表项中的计数器加1;所述第一IP报文携带的信息包括所述第一IP报文的源IP地址和目的IP地址,或者所述第一IP报文的源IP地址、目的IP地址,源端口号和目的端口号。

[0058] 本申请实施例的第四方面公开了一种网络设备,包括:存储器,以及与所述存储器通信的处理器;

[0059] 所述存储器,用于存储处理报文的程序代码;

[0060] 所述处理器,用于执行所述存储器保存的程序代码,以实现上述第一方面以及各种可能的设计中的操作。

[0061] 本申请实施例的第五方面公开了一种终端设备,用作第一终端设备,若所述第一终端设备发送的数据流支持反向拥塞探测机制,所述第一终端设备包括:

[0062] 发送单元,用于向网络设备发送第一因特网协议IP报文,所述第一IP报文的IP报文头中包括反向显式拥塞通告RECN标记,所述RECN标记用于标识所述第一IP报文所属的数据流支持反向拥塞探测;

[0063] 接收单元,用于接收来自所述网络设备的第一标记报文,所述第一标记报文包括拥塞发生CE标记,所述CE标记用于标识所述网络设备在传输报文时发生了拥塞;

[0064] 处理单元,用于基于所述第一标记报文执行拥塞响应。

[0065] 在一种可能的设计中,若所述第一终端设备发送的数据流支持反向拥塞探测机制和正向拥塞探测机制;

[0066] 所述接收单元,还用于接收来自所述网络设备的第二标记报文,所述第二标记报文包括CE标记;

[0067] 所述处理单元,还用于确定所述第二标记报文的源端口号和目的端口号,并基于所述第二标记报文的源端口号和目的端口号确定所述第二标记报文所属的数据流;若确定所述数据流所支持的拥塞探测机制为反向拥塞探测机制,则所述第一终端设备执行对所述数据流的拥塞处理;若确定所述数据流所支持的拥塞探测机制为正向拥塞探测机制,则向所述网络设备发送响应ACK消息,所述ACK消息用于指示第二终端设备执行对所述数据流的拥塞处理,所述ACK消息中包括CE标记。

[0068] 在本申请实施例的第六方面公开了一种终端设备,用作第一终端设备,所述第一终端设备发送的数据流支持反向拥塞探测机制,或者,所述第一终端设备发送的数据流可以支持反向拥塞探测机制,也可以支持正向拥塞探测机制,所述第一终端设备包括:存储器,以及与所述存储器通信的处理器;

[0069] 所述存储器,用于存储处理报文的程序代码;

[0070] 所述处理器,用于执行所述存储器保存的程序代码,以实现上述第二方面以及各种可能的设计中的操作。

[0071] 本申请实施例的第七方面公开了一种网络系统,包括:第一终端设备和网络设备;

- [0072] 所述网络设备为上述第三方面或第四方面所述的网络设备；
- [0073] 所述第一终端设备可以为上述第五方面或第六方面所述的终端设备。
- [0074] 本申请实施例的第八方面提供了一种计算机可读存储介质，用于存储计算机程序，该计算机程序包括用于执行第一方面、第二方面、第一方面的任一可能的设计或第二方面的任一可能的设计中的方法的指令。

附图说明

- [0075] 为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据提供的附图获得其他的附图。
- [0076] 图1为本申请实施例公开的一种端到端的网络连接示意图；
- [0077] 图2为本申请实施例公开的网络设备的结构示意图；
- [0078] 图3为本申请实施例公开的一种报文处理方法的流程示意图；
- [0079] 图4为本申请实施例公开的另一种报文处理方法的流程示意图；
- [0080] 图5为本申请实施例公开的另一种报文处理方法的流程示意图；
- [0081] 图6为本申请实施例公开的另一种报文处理方法的流程示意图；
- [0082] 图7为本申请实施例公开的一种网络设备的结构示意图；
- [0083] 图8为本申请实施例公开的一种第一终端设备的结构示意图；
- [0084] 图9为本申请实施例公开的另一种网络设备的结构示意图；
- [0085] 图10为本申请实施例公开的另一种第一终端设备的结构示意图；
- [0086] 图11为本申请实施例公开的一种网络系统的结构示意图。

具体实施方式

[0087] 本申请实施例提供了一种报文处理方法、网络设备和系统，用于缩短拥塞探测时间，降低网络拥塞控制带来的时延，并为传输短寿命数据流和对传输时延敏感的业务提供有利条件。

[0088] 本申请实施例和权利要求书及附图中的术语“第一”、“第二”、“第三”和“第四”等是用于区别不同对象，而不是用于描述特定顺序。此外，术语“包括”和“具有”不是排他的。例如，包括了一系列步骤或单元的过程、方法、系统、产品或设备没有限定于已列出的步骤或单元，还可以包括没有列出的步骤或单元。

[0089] 本申请实施例公开的报文处理方法通过以下具体实施例进行详细说明。

[0090] 如图1所示，为本申请实施例公开的一种端到端的网络连接示意图。该示意图中包括：网络设备10，发送端设备11和接收端设备12。发送端设备11和接收端设备12通过网络设备10进行报文交互。

[0091] 本申请实施例所公开的报文处理方法应用于因特网协议（英文：Internet Protocol, IP）层，发送端设备11向接收端设备发送的报文封装为IP报文格式。可选的，在该IP报文头中标记拥塞探测中所涉及的信息。IP报文头的部分格式如表1所示。

[0092] 表1：IP报文头部分格式

	0	1	2	3	4	5	6	7
[0093]								ECN FIELD

[0094] 显式拥塞通告字段(英文:Explicit Congestion Notification FIELD,ECN FIELD),占用第6bit位和第7bit位,利用不同的置位标记拥塞探测中的不同信息。如表2所示,为针对ECN FIELD所进行的置位,以及置位所对应的支持拥塞探测的机制。

[0095] 表2:ECN FIELD置位及其标识含义

[0096]

第6bit位	第7bit位	标识含义
0	0	Not-ECT
0	1	ECT (1)
1	0	ECT (0)
1	1	CE

[0097] 在表2中,“00”:用于标识当报文在网络设备中发生拥塞时,不支持ECN,即Not-ECT。

[0098] “01”和“10”:当发送端设备,网络设备和接收端设备之间仅支持反向拥塞探测机制,即反向显式拥塞通告(英文:Reverse Explicit Congestion Notification,RECN)机制时,“01”和“10”均用于标识当报文在传输过程中发生拥塞时,支持RECN机制,即ECT (0) 和ECT (1) 都表示支持RECN机制。

[0099] 在发送端设备,网络设备和接收端设备之间即支持RECN机制,也支持正向ECN机制的情况下,当置位为“01”,即ECT (1)时,用于标识报文在传输过程中发生拥塞时支持RECN机制。当置位为“10”,即ECT (0)时,用于标识报文在传输过程中发生拥塞时支持正向ECN机制。当然本申请实施例也并不仅限于此,在支持多种探测机制的情况下,ECN FIELD所进行的置位只要能够区分RECN机制和正向ECN机制即可。

[0100] “11”:用于标识网络设备在传输报文时发生了拥塞(英文:Congestion Experienced,CE)。

[0101] 作为举例,如图2所示,为本申请实施例公开的网络设备10的结构示意图。该网络设备10包括:上行单板201,交换网板202和下行单板203。

[0102] 上行单板201包括:上行物理接口卡(英文:Physical Interface Card,PIC),与上行PIC连接的上行包转发引擎(英文:Packet Forwarding Engines,PFE),与上行PFE连接的上行流量管理器(英文:Traffic Manager,TM),与上行TM连接的上行转发信息表(英文:Forwarding Information Base,FIC),该上行FIC与交换网板202连接。

[0103] 下行单板203包括:与交换网板202连接的下行FIC,与下行FIC连接的下行TM,与下行TM连接的下行PFE,与下行PFE连接的下行PIC。

[0104] 需要说明的是,本发明实施例中的上行单板和下行单板是针对某个数据报文在网络设备内部转发流程的角度来说的。上行单板(即转发单板)是该数据报文进入该网络设备的接口(即入接口)所在的单板,通过上行单板的转发处理将该报文发送到交换网板(也叫交换网卡)。通过交换网板的交换,将该数据报文发送到下行单板,该下行单板执行转发处理,通过出接口把该数据报文从该网络设备上发送出去。故下行单板是该数据报文离开该网络设备的接口(即出接口)所在的单板。由此可以看出,上行是指从网络设备接收一个数

据报文向交换网板转发的流向,下行是指从交换网板到将一个数据报文从网络设备上转发出去的流向。应理解,同一块单板针对于一个数据报文来说,可以是上行单板;而针对另一个数据报文来说,也可以是下行单板。还应理解,当一个数据报文经过一个网络设备时的入接口和出接口在同一个单板上时,针对该数据报文来说,上行单板和下行单板是指同一块单板。

[0105] 本申请实施例中,发送端设备11经过网络设备10向接收端设备12发送报文。当报文在网络设备10中发生拥塞,由网络设备10对报文设置CE标记,并直接反馈给发送端设备,相较于现有技术,不需要经过接收端设备,从而缩短了至少一半的拥塞探测时间。

[0106] 在本申请实施例中,网络设备可以是路由器,也可以是交换机等用于转发处理的设备。

[0107] 如图3所示,为本申请实施例公开的一种报文处理方法的流程示意图,在本申请实施例中,第一终端设备基于TCP建立与第二终端设备之间的连接。第一终端设备与第二终端设备之间的数据流只支持反向拥塞探测机制。该报文处理方法包括:

[0108] S301:第一终端设备经过网络设备向第二终端设备发送数据流中的IP报文。

[0109] 作为举例,该第一终端设备,可以为本申请实施例图1中示出的发送端设备。该网络设备,可以为本申请实施例图1中示出的网络设备。该第二终端设备,可以为本申请实施例图1中示出的接收端设备。

[0110] 该数据流由多个IP报文构成。以接收到数据流中的一个IP报文为例进行说明。

[0111] S302:网络设备在检测到网络设备发生拥塞时,网络设备获得来自第一终端设备的第一IP报文的源IP地址和目的IP地址,将该第一IP报文的源IP地址存储于第一数据表项的目的IP地址中,将该第一IP报文的源IP地址存储于第一数据表项的源IP地址中。

[0112] 作为举例,网络设备通过检测存储队列中的存储容量是否大于预设阈值,确定网络设备是否发生拥塞。当检测到存储队列中的存储容量大于预设阈值时,确定网络设备发生拥塞。该预设阈值通常小于存储队列的总容量,该网络设备的存储队列的总容量与网络设备本身有关,不同的网络设备存储队列的总容量不同。存储队列的容量不仅可以以包的个数来表示,也可以以KB大小来表示。

[0113] 网络设备在检测到发生了拥塞后,按照预设规则确定发生拥塞的IP报文。该预设规则与网络设备采用什么方式对接收到的报文进行存储有关。例如:网络设备采用一个存储队列的方式存储接收到的不同IP报文,也可以采用分类的方式,将接收到的IP报文按照所属数据流的不同,进行分类存储于对应的存储队列中。

[0114] 在网络设备检测到存储队列中的存储容量大于预设阈值时,基于该预设规则选择当前发生拥塞的报文。该预设规则包括主动队列管理(英文:Active Queue Management, AQM)中的各种方法。作为举例,该预设规则包括:随机早期检测(random early detection, RED),延迟控制(英文:Controlled Delay, Co DEL)和比例积分增强(英文:Proportional Integral Enhanced, PIE),即队尾丢弃算法。

[0115] 该第一数据表项属于网络设备的数据表中的一条表项,该数据表的每一列对应一个项,每一行对应一个数据表项。该数据表可以由网络设备预先建立,也可以在使用时再建立。同样的,该数据表中的数据表项也可以由网络设备预先建立,也可以在网络设备获得来自第一终端设备的第一IP报文的源IP地址和目的IP地址时,再进行建立。

[0116] 该数据表如表3所示,项的类型包括:报文的源IP(英文:Source IP,SIP)地址和目的IP(英文:Destination,DIP)地址。其中,还可以设置表项编号(ID)。

[0117] 作为举例,第一行为第一数据表项。

[0118] 表3

[0119]

ID	SIP	DIP
1	第一IP报文的源IP地址	第一IP报文的源IP地址

[0120] 网络设备为了后续采用反向拥塞探测机制进行拥塞探测。网络设备获取第一IP报文的源IP地址和目的IP地址,然后在存储于表3示出的第一数据表项时,将获取的第一IP报文的源IP地址和目的IP报文进行交换后存储,以便于后续查询反向传输的IP报文进行标记。

[0121] S303:网络设备获得来自第二终端设备的第二IP报文,并确定第二IP报文的源IP地址与第一数据表项中的SIP地址相同,第二IP报文的源IP地址与第一数据表项中的DIP地址相同,对第二IP报文设置CE标记,得到第一标记报文。

[0122] 该CE标记用于标识网络设备在传输报文时发生了拥塞。

[0123] S304:网络设备向第一终端设备发送第一标记报文。

[0124] S305:第一终端设备获得来自网络设备的第一标记报文,基于该第一标记报文执行拥塞响应。

[0125] 拥塞响应有多种方式,在本申请实施例中,第一终端设备、第二终端设备和网络设备之间仅支持反向拥塞探测机制。因此,该第一终端设备基于该第一标记报文,确定自身发送的IP报文在网络设备处发生了拥塞。由此,第一终端设备对向网络设备发送的数据流进行拥塞处理,该拥塞处理包括:降低数据流中IP报文的发送速率,或者对传输的数据流进行丢包处理,或者,停止某一数据流的传输。对于该第一终端设备对数据流所执行的拥塞处理,本申请实施例并不进行具体限定,所执行的拥塞处理只要可以缓解网络设备的拥塞状态即可。

[0126] 基于上述本申请实施例公开的以第一IP报文进行举例说明的报文处理方法,在具体应用中,网络设备在确定发生拥塞后,可选的,可以对发生拥塞后接收到的所有IP报文进行处理,也可以对部分IP报文进行处理。

[0127] 在具体实现过程中,该网络设备获取存储容量大于预设阈值之后接收到的IP报文进行处理。作为举例,若该网络设备的存储队列的总容量为最大存储100个报文,该预设阈值为50个IP报文。在存储队列中所存储的报文超过50个时,从第51个IP报文开始进行处理。

[0128] 因此,结合表3,在数据表中的项还包括计数器,如表4所示。网络设备每接收到一个与第一IP报文的源IP地址和目的IP地址分别相同的IP报文时,将计数器加1;网络设备每接收到一个与第二IP报文的源IP地址和目的IP地址分别相同的IP报文时,将计数器减1。

[0129] 最终,该计数器中的计数用于指示待设置CE标记的IP报文的个数,该待设置CE标记的IP报文的源IP地址与第二IP报文的源IP地址相同,待设置CE标记的IP报文的源IP地址与第二IP报文的源IP地址相同,待设置CE标记的IP报文的源IP地址与第二IP报文的源IP地址相同。

[0130] 表4

[0131]

ID	SIP	DIP	Num
1	第一IP报文的目的IP地址	第一IP报文的源IP地址	Num

[0132] 在该第一数据表中，SIP中存储的是发生了拥塞的IP报文的目的IP地址，作为后续需要设置CE标记的IP报文的源IP地址。

[0133] DIP中存储的是发生了拥塞的IP报文的源IP地址，作为后续需要设置CE标记的IP报文的目的IP地址。

[0134] 个数(英文: Number, Num)，用于指示计数器，该Num中存储的是需要设置CE标记的IP报文的个数。该Num中的数值是动态变化的，网络设备在存储容量大于预设阈值之后每获取一个IP报文，则对该Num表项中的数值进行加1。网络设备每设置一个CE标记，则对该Num表项中的数值进行减1。

[0135] 作为举例，该网络设备可以为本申请实施例图1中示出的网络设备10。则网络设备在本申请实施例图2中示出的下行PFE中获取的来自第二终端设备的IP报文中的源IP地址和目的IP地址，将该IP报文中的目的IP地址存储于SIP地址中，将该IP报文中的源IP地址存储于DIP地址中，并同时更新Num。

[0136] 基于上述本申请实施例公开的报文处理方法，结合表4示出的数据表，若第一数据表项中还包括计时器(英文: Timer)，具体如表5所示。

[0137] 以网络设备对获得的来自第一终端设备的第一IP报文进行处理，以及对网络设备获得来自第二终端设备的第二IP报文设置CE标记，得到第一标记报文的过程为例进行说明，如图4所示，该报文处理方法包括：

[0138] S401: 网络设备获得来自第一终端设备的第一IP报文的源IP地址和目的IP地址，将第一IP报文中的目的IP地址存储于第一数据表项的SIP地址中，将第一IP报文中的源IP地址存储于第一数据表项的DIP地址中，并将第一数据表项的Num加1，第一数据表项的Timer重置计时。

[0139] 表5

ID	SIP	DIP	Num	Timer
[0140] 1	第一IP报文的 目的IP地址	第一IP报文的 源IP地址	Num	一个报文的 计时

[0141] 网络设备在表4示出的数据表中增加一项Timer，其他项相同，可以参见表4中的记载，这里不再进行赘述。在此，仅对第一数据表项中的Timer进行详细说明。

[0142] Timer中对应的值为针对一个报文的计时。

[0143] Timer为动态变化的。网络设备在生成第一数据表项或对第一数据表项中的Num减1时，该Timer重置计时。可选的，网络设备可以设置该Timer中的最大计时为1个RTT。

[0144] 可选的，网络设备可以采用正向计时方式，也可以采用倒计时的方式。

[0145] S402: 在计时器计时过程中，网络设备获得来自第二终端设备的第二IP报文，该第一数据表项中的Num数值减1，Timer重置计时。

[0146] 该第二IP报文的源IP地址与第一数据表项中的SIP地址相同，第二IP报文的目的IP地址与第一数据表项中的DIP地址相同。

[0147] S403: 网络设备对第二IP报文设置CE标记，得到第一标记报文。

[0148] 网络设备执行S403所依据的原理以及具体执行过程与本申请实施例中图3公开的S304记载的内容相同,这里不再进行赘述。

[0149] S404:在Timer计时过程中,若Timer的计时达到预设计时,网络设备删除第一数据表项。

[0150] 在Timer计时过程中,若网络设备一直未获得来自第二终端设备的第二IP报文,或者源IP地址与第一数据表项中的SIP地址相同,目的IP地址与第一数据表项中的DIP地址相同的IP报文,则Timer计时在达到预设计时,网络设备将删除第一数据表项。

[0151] 可选的,在计时器计时的过程中,若网络设备采用正计时方式,则预设计时等于预先设置的该Timer表项中的计时器的最大计时;若网络设备采用倒计时方式,则预设计时则为0。

[0152] 作为举例,预先设置Timer的计时为0.25秒。若网络设备采用正向计时方式,则在重置计时后从0开始递增计时,当计时达到0.25秒或超过0.25秒时,则网络设备删除第一数据表项;若网络设备采用倒计时的方式,则在重置计时后从0.25秒开始递减计时,当计时达到0,则网络设备删除第一数据表项。

[0153] 基于上述本申请实施例公开的报文处理方法,结合表3示出的数据表,若第一数据表项中还包括源端口(SPort)号和目的端口(DPort)号,具体如表6所示。

[0154] 以网络设备对获得的来自第一终端设备的第一IP报文进行处理,以及对网络设备获得来自第二终端设备的第二IP报文设置CE标记,得到第一标记报文的过程为例进行说明,如图5所示,该报文处理方法包括:

[0155] S501:网络设备获取来自第一终端设备发送的第一IP报文中的源IP地址,目的IP地址,源端口号和目的端口号,将第一IP报文的源IP地址存储于第一数据表项的SIP地址中,将第一IP报文的源IP地址存储于第一数据表项的DIP地址中,将第一IP报文的源端口号存储于第一数据表项的DPort中,将第一IP报文的源端口号存储于第一数据表项的SPort中。

[0156] 在本申请实施例中,相较于表3,该第一数据表项中增加了DPort和SPort,其他与本申请实施例公开的表3相同,可以参见上述记载,这里不再进行赘述。

[0157] 在本申请实施例中,网络设备获得到第一终端设备的第一IP报文的源端口号和目的端口号进行存储,可以在后续查找需要设置CE标记的反向报文时,确定设置CE标记的报文与第一IP报文属于同一个数据流。

[0158] 表6

[0159]

ID	SIP	DIP	SPort	DPort
1	第一IP报文的源IP地址	第一IP报文的源IP地址	第一IP报文的源端口号	第一IP报文的源端口号

[0160] 若第一数据表项中还包括Num项,则网络设备在执行S501的过程中,对第一数据表项的Num中的数值加1。即网络设备每接收到一个与第一IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同的IP报文时,将Num中的数值加1。

[0161] 若第一数据表项中还包括Timer项,则网络设备在执行S501的过程中,Timer进行计时。

[0162] S502:网络设备获得来自第二终端设备的第二IP报文,并确定第二IP报文的源IP地址与第一数据表项中的SIP地址相同,第二IP报文的源IP地址与第一数据表项中的DIP地址相同,第二IP报文的源端口号与第一数据表项中的SPort相同,第二IP报文的源端口号与第一数据表项中的DPort相同,第二IP报文设置CE标记,得到第一标记报文。

[0163] 本申请实施例中,该S502对第二IP报文设置CE标记的过程与图3所对应的实施例中S304所记载的过程相同,这里不再进行赘述。

[0164] 若第一数据表项中还包括Num项,则网络设备在执行S502的过程中,对第一数据表项的Num中的数值减1。即网络设备每接收到一个与第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同的IP报文时,将Num中的数值减1。

[0165] S503:网络设备向第一终端设备发送第一标记报文。

[0166] S504:第一终端设备接收该第一标记报文,并依据该第一标记报文执行拥塞响应。

[0167] 本申请实施例中,该S504中第一终端设备依据第一标记报文执行拥塞响应的过程,与图3所对应的实施例中S306所记载的过程相同,这里不再进行赘述。

[0168] 基于上述本申请实施例公开的报文处理方法,可选的,可以对需要进行处理的IP报文添加标识信息。该标识信息的形式并不进行具体限定,可以是任意形式。

[0169] 在具体实现过程中,网络设备在获得来自第一终端设备的IP报文之后,对该IP报文添加标识信息,该标识信息用于指示网络设备根据IP报文携带的信息生成第一数据表项和/或对第一数据表项中的计数器加1。可选的,该IP报文携带的信息包括IP报文的源IP地址和目的IP地址;或者,该IP报文携带的信息包括IP报文的源IP地址、目的IP地址、源端口号和目的端口号。

[0170] 以网络设备获得的来自第一终端设备的第一IP报文为例。网络设备在获得来自第一终端设备的IP报文之后,对该第一IP报文添加标识信息,该标识信息用于指示网络设备根据第一IP报文携带的信息生成第一数据表项和/或对第一数据表项中的计数器加1。可选的,该第一IP报文携带的信息包括第一IP报文的源IP地址和目的IP地址;或者,该第一IP报文携带的信息包括第一IP报文的源IP地址、目的IP地址、源端口号和目的端口号。

[0171] 作为举例,该网络设备可以为本申请实施例图1中示出的网络设备10,则网络设备在本申请实施例图2中示出的上行TM中,对依据预设规则确定存储容量大于预设阈值之后获得的来自第一终端设备的IP报文,添加标识信息。

[0172] 基于上述本申请实施例公开的报文处理方法,在本申请实施例中,第一终端设备和网络设备之间所支持的反向拥塞探测机制,可以在第一终端设备和网络设备之间预先进行设定,也可以在第一终端设备和网络设备之间传输的报文中添加用于指示反向拥塞探测机制的RECN标记,本申请实施例对此并不进行限定。该RECN标记可以采用新增字段的方式,设置于报文中,也可以在特定的字段中进行添加,本申请实施例对此并不进行限定。

[0173] 可选的,第一终端设备向网络设备发送的第一IP报文的报文头中包括RECN标记,该RECN标记用于标识该第一IP报文所属的数据流支持反向拥塞探测机制。网络设备获得来自第一终端设备的第一IP报文,确定该第一IP报文的报文头中ECN FIELD的置位信息。

[0174] 作为举例,若第一IP报文的报文头中ECN FIELD的置位信息为“01”,则确定该第一ECN标记为ECT(1),或者,若第一IP报文的报文头中的ECN FIELD字段的置位信息为“10”,则确定该第一ECN标记为ECT(0)。在本申请实施例中,ECT(0)和ECT(1)均用于标识当该第一IP

报文在网络设备中发生拥塞时,支持RECN进行拥塞探测。

[0175] 可选的,在网络设备对获得的来自第二终端设备的第二IP报文设置CE标记时,该CE标记可以采用新增字段的方式,设置于第二IP报文中。若第二IP报文的报文头中存在ECN FIELD字段,也可以直接将已有ECN FIELD字段中的ECT (1) 或ECT (0) 中的“0”位置1。即将ECN FIELD字段置位为“11”,得到CE标记。

[0176] 基于上述本申请实施例公开的报文处理方法,网络设备在接收到终端设备发送的IP报文,并基于网络设备和终端设备之间所支持的反向拥塞探测机制,通过在反向传输的IP报文中设置CE标记,然后反馈该终端设备,可以使终端设备直接通过该CE标记判断自身的报文传输发生拥塞。从而及时做出拥塞响应,降低该报文的发送速率。相较于现有技术而言,本申请实施例公开的报文处理方法,在拥塞探测时,不需要发送给对端设备进行响应,也不需要依赖终端设备自身判断对应该报文的计时是否超时,减少了拥塞探测的时间,进而降低时延。

[0177] 如图6所示,为本申请实施例公开的另一种报文处理方法的流程示意图,在本申请实施例中,第一终端设备基于TCP建立与第二终端设备之间的连接。且第一终端设备与第二终端设备之间发送不同数据流时,确定各个数据流所支持的不同类型拥塞探测机制。不同类型拥塞探测机制指发送不同类型的拥塞通告,该拥塞通告类型包括RECN和正向ECN。该报文处理方法包括:

[0178] S601:第一终端设备经过网络设备向第二终端设备发送第三IP报文。

[0179] 作为举例,该第一终端设备,可以为本申请实施例图1中示出的发送端设备。该网络设备,可以为本申请实施例图1中示出的网络设备。该第二终端设备,可以为本申请实施例图1中示出的接收端设备。

[0180] 作为举例,在本申请实施例中,不同类型拥塞探测机制为:

[0181] 该第三IP报文的报文头中ECN FIELD的置位为“00”,即Not-ECT。该Not-ECT用于标识当该第三IP报文在网络设备中发生拥塞时,不支持ECT机制进行拥塞探测。

[0182] 该第三IP报文的报文头中ECN FIELD的置位为“01”,即ECT (1)。该ECT (1) 用于标识当该第三IP报文在网络设备中发生拥塞时,支持RECT机制进行拥塞探测。

[0183] 该第三IP报文的报文头中ECN FIELD的置位为“10”,即ECT (0)。该ECT (0) 用于标识当该第三IP报文在网络设备中发生拥塞时,支持正向ECN机制进行拥塞探测。

[0184] 该第三IP报文的报文头中ECN FIELD的置位为“11”,即CE。该CE用于标识报文在传输过程中发生拥塞。

[0185] S602:网络设备获得来自第一终端设备的第三IP报文,确定该第三IP报文的报文头中的ECN FIELD的置位信息,若该置位信息标识该第三IP报文的报文头中包含第一ECN标记,则执行S603-S604;若该置位信息标识该第三IP报文的报文头中包含第二ECN标记,则执行S605-S607;若该置位信息标识该第三IP报文的报文头中包含Not-ECT标记,则执行S608-S609。

[0186] S603:网络设备在检测网络设备发生拥塞时,依据预设规则确定发生拥塞的为第三IP报文,网络设备确定该第三IP报文头中包含的第一ECN标记为ECT (1),则该网络设备对获取到的第三IP报文设置CE标记,得到第二标记报文,并将该第二标记报文发送给第一终端设备。

[0187] 基于设备之间支持不同类型拥塞探测机制的置位,网络设备确定接收到的第三IP报文头中的置位为“01”,则确定执行RECN机制进行拥塞探测。作为举例,网络设备可以采用上述本申请实施例图5中示出的对第二IP报文设置CE标记的方式,对第三IP报文设置CE标记,得到第二标记报文。并将该第二标记报文发送给第一终端设备。

[0188] S604:第一终端设备获得来自网络设备的第二标记报文,确定该第二标记报文所属于的数据流,该第一终端设备基于该确定的数据流所支持的拥塞探测机制,执行相应地拥塞处理。

[0189] 该第二标记报文包括CE标记,该CE标记用于标识报文传输中发生了拥塞,该第二标记报文包括源端口号和目的端口号,该源端口号和目的端口号用于标识该第二标记报文所属的数据流。

[0190] 在第一终端设备接收到包括CE标记的报文后,可以通过该报文的源端口号和目的端口号确定该报文所属于的数据流。第一终端设备基于该确定的数据流所支持的拥塞探测机制,在第一终端设备对该数据流执行相应的相应地拥塞处理。

[0191] 可选的,当第一终端设备确定该数据流支持RECN,则第一终端设备确定自己发送的报文在网络设备中发生拥塞,执行拥塞响应,降低发送该数据流中的报文的发送速率,实现拥塞控制。

[0192] S605:网络设备在检测到网络设备发生拥塞时,依据预设规则确定发生拥塞的为第三IP报文,网络设备确定第三IP报文头中包含的第二ECN标记为ECT(0),则该网络设备对获取到的第三IP报文设置CE标记,得到第三标记报文,并将该第三标记报文发送给第一终端设备。

[0193] 基于设备之间支持不同类型拥塞探测机制的置位,网络设备确定接收到的第三IP报文头中的置位为“10”,则确定执行正向ECN机制进行拥塞探测。具体为:将第三IP报文中的置位“10”更改为“11”,然后发送给第二终端设备。

[0194] S606:第二终端设备获得来自网络设备的第三标记报文,向第一终端设备发送给在TCP头部添加拥塞标记的ACK消息,直至接收到第一终端设备发送的确定消息后停止向第一终端设备发送具有拥塞标记的ACK消息。

[0195] 第二终端设备在接收到设置CE标记的第三标记报文时,可以根据该第三标记报文包括的源端口号和目的端口号,确定该第三标记报文所属的数据流。作为举例,确定该数据流所支持的拥塞通告类型为正向ECN机制,则确定对端发生拥塞。也就是说,第一终端设备发生拥塞。

[0196] S607:第一终端设备接收到携带有拥塞标记的ACK消息,向该第二终端设备发送确定消息,并进行相应地拥塞响应。

[0197] S608:网络设备确定第三IP报文头中包含Not-ECT标记,则在检测到网络设备发生拥塞时,依据预设规则进行丢包处理。

[0198] 基于设备之间支持不同类型拥塞探测机制的置位,网络设备确定获得的第三IP报文头中的置位为“00”,则确定该数据流不支持ECT机制进行拥塞探测。因此,网络设备执行丢包处理。此时,第二终端设备将不会接收到已经被丢弃的报文,也不会向第一终端设备发送ACK消息。第一终端设备在传输的第三IP报文对应的计时器超时后,即判断报文丢失,确定网络发生拥塞,则进行拥塞响应。

[0199] 可选的,若第一终端设备为本申请实施例图1中示出的接收端设备,则在接收到网络设备发送的标记为CE的报文时,可以根据该标记为CE的报文中包括的源端口号和目的端口号,确定该标记为CE的报文所属的数据流,以及确定该数据流所支持的拥塞通告类型。若该数据流支持正向ECN机制,则确定对端发生拥塞。若该数据流支持RECT机制,则确定自己发生拥塞。

[0200] 在本申请实施例中,因第一终端设备与第二终端设备之间发送的不同数据流,支持不同类型拥塞探测机制。在本申请实施例中作为发送端设备的第一终端设备在发送报文时,基于在报文头中ECN FIELD的不同置位,区分各个数据流所支持的拥塞探测机制。网络设备根据接收到的报文中的报文头的ECN FIELD不同置位,执行相应地拥塞探测机制。基于不同数据流支持不同拥塞探测机制的方式,使网络设备在执行拥塞探测时,能够更加灵活。

[0201] 基于上述本申请实施例一公开的报文处理方法,本申请实施例还对应公开了执行该报文处理方法的网络设备。

[0202] 如图7所示,本申请实施例公开的网络设备70包括:

[0203] 处理单元701,用于若网络设备发生拥塞时,获得来自第一终端设备的第一IP报文的源IP地址和目的IP地址,将第一IP报文的源IP地址存储于第一数据表项的目的IP地址中,将第一IP报文的源IP地址存储于第一数据表项的源IP地址中,以及获得来自第二终端设备的第二IP报文,第二IP报文的源IP地址与第一数据表项中的源IP地址相同,第二IP报文的源IP地址与第一数据表项中的源IP地址相同,对第二IP报文设置拥塞发生CE标记,得到第一标记报文,CE标记用于标识所述网络设备在传输报文时发生了拥塞。

[0204] 发送单元702,用于向第一终端设备发送第一标记报文。

[0205] 若处理单元701所建立的第一数据表项中还包括计数器,该计数器用于指示待设置CE标记的IP报文的个数,所述待设置CE标记的IP报文的源IP地址和目的IP地址与所述第二IP报文的源IP地址和目的IP地址分别相同。

[0206] 相应地,该处理单元701,用于每接收到一个与第一IP报文的源IP地址和目的IP地址相同的IP报文时,将计数器加1,以及每得到一个与第二报文的源IP地址和目的IP地址相同的IP报文时,将计数器减1。最终得到待设置CE标记的IP报文的个数。

[0207] 若处理单元701所建立的第一数据表项中还包括计时器。

[0208] 相应地,该处理单元701,还用于在生成第一数据表项或对第一数据表项中的计数器减1时,计时器重置计时,若计时器的计时达到预设计时,删除所述第一数据表项。

[0209] 若处理单元701所建立的第一数据表项中还包括源端口号和目的端口号;

[0210] 相应地,处理单元701,还用于获取来自第一终端设备的第一IP报文中的源端口号和目的端口号,将第一IP报文的源端口号存储于第一数据表项的目的端口号中,将第一IP报文的源端口号存储于第一数据表项的源端口号中;以及确定第二IP报文的源IP地址与第一数据表项中的源IP地址相同,第二IP报文的源IP地址与第一数据表项中的源IP地址相同,确定第二IP报文的源端口号与第一数据表项中的源端口号相同,第二IP报文的源端口号与第一数据表项中的源端口号相同,对第二IP报文设置拥塞发生CE标记。

[0211] 若待设置CE标记的IP报文的源IP地址、目的IP地址、源端口号和目的端口号与第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同;

[0212] 相应地,处理单元701,用于每接收到一个与第一IP报文的源IP地址、目的IP地址、

源端口号和目的端口号分别相同的IP报文时,将计数器加1;每接收到一个与第二IP报文的源IP地址、目的IP地址、源端口号和目的端口号分别相同的IP报文时,将计数器减1。

[0213] 基于上述公开的网络设备,该处理单元701,还用于对获得第一IP报文之后,对该第一IP报文添加标识信息,该标识信息用于指示网络设备根据第一IP报文携带的信息生成第一数据表项和/或对第一数据表项中的计数器加1。所获得的第一IP报文携带的信息包括第一IP报文的源IP地址和目的IP地址,或者所述第一IP报文的源IP地址、目的IP地址,源端口号和目的端口号。

[0214] 以上本申请实施例公开的网络设备中的各个单元的执行过程,可以参见本申请实施例上述报文处理方法中涉及网络设备的记载,原理和执行方式相同,这里不再进行赘述。

[0215] 基于上述本申请实施例一公开的报文处理方法,本申请实施例还对应公开了执行该报文处理方法的第一终端设备。

[0216] 如图8所示,为本申请实施例示出的该第一终端设备80的结构示意图。

[0217] 该第一终端设备80包括:发送单元801,接收单元802和处理单元803。

[0218] 发送单元801,用于向网络设备发送第一IP报文,第一IP报文的IP报文头中包括RECN标记,RECN标记用于标识第一IP报文所属的数据流支持反向拥塞探测。

[0219] 若该第一终端设备80发送的数据流支持反向拥塞探测机制:

[0220] 接收单元802,用于接收来自网络设备的第一标记报文,第一标记报文包括拥塞发生CE标记,CE标记用于标识网络设备在传输报文时发生了拥塞。

[0221] 处理单元803,用于基于第一标记报文执行拥塞响应。

[0222] 若第一终端设备80发送的数据流支持反向拥塞探测机制和正向拥塞探测机制:

[0223] 接收单元802,还用于接收来自网络设备的第二标记报文,第二标记报文包括CE标记;

[0224] 处理单元803,还用于确定第二标记报文的源端口号和目的端口号,并基于第二标记报文的源端口号和目的端口号确定第二标记报文所属的数据流;若确定数据流所支持的拥塞探测机制为反向拥塞探测机制,则第一终端设备执行对数据流的拥塞处理;若确定数据流所支持的拥塞探测机制为正向拥塞探测机制,则向网络设备发送响应ACK消息,ACK消息用于指示第二终端设备执行对数据流的拥塞处理,ACK消息中包括CE标记。

[0225] 以上本申请实施例公开的第一终端设备中的各个单元的执行过程,可以参见本申请实施例上述报文处理方法中涉及第一终端设备的记载,原理和执行方式相同,这里不再进行赘述。

[0226] 结合本申请实施例公开的报文处理方法,本申请实施例所公开的网络设备和第一终端设备也可以直接用硬件、处理器执行的存储器,或者二者的结合来实施。

[0227] 因此,本申请还对应上述本申请实施例公开的报文处理方法,公开了如图9所示的网络设备90,如图10所示的第一终端设备100。

[0228] 如图9所示,该网络设备90包括:存储器901、总线902、处理器903和接口904。其中,接口904可以通过无线或有线的方式实现,具体来讲可以是例如网卡等元件,上述存储器901、处理器903和接口904通过总线902连接。

[0229] 该接口904具体可以包括发送器和接收器,用于网络设备与上述本申请实施例中的第一终端设备和第二终端设备之间收发报文。具体过程可参见上述本申请实施例涉及网

络设备收发报文的相应部分,这里不再赘述。

[0230] 所述处理器903用于执行上述申请实施例涉及网络设备执行对第一IP报文和第二IP报文进行处理的部分。具体过程可参见,这里不再赘述。

[0231] 存储器901包括操作系统和应用程序,用于存储报文处理的操作程序、代码或指令,当处理器903或硬件设备在执行报文处理时,调用并执行这些程序、代码或指令时可以完成图3-图7中涉及网络设备对报文进行处理的过程。具体过程可参见上述本申请实施例相应地部分,这里不再赘述。

[0232] 可以理解的是,图9仅仅示出了网络设备的简化设计。在实际应用中,网络设备可以包含任意数量的接口,处理器,存储器等,而所有可以实现本申请的网络设备都在本申请的保护范围之内。

[0233] 另外,本申请实施例提供了一种计算机存储介质,用于储存为上述网络设备所用的计算机软件指令,其包含用于执行上述图3-图7所示实施例所设计的程序。

[0234] 图10为本申请实施例公开的第一终端设备100的硬件结构示意图。图10所示的第一终端设备100可以执行上述本申请实施例中第一终端设备所执行的相应步骤。

[0235] 如图10所示,该第一终端设备100包括:存储器1001、总线1002、处理器1003和接口1004。其中,接口1004可以通过无线或有线的实现,具体来讲可以是例如网卡等元件,上述存储器1001、处理器1003和接口1004通过总线1002连接。

[0236] 该接口1004具体可以包括发送器和接收器,用于第一终端设备与上述本申请实施例中的网络设备之间收发报文。例如,该接口1004的具体过程可参见上述本申请实施例中,第一终端设备与网络设备之间收发报文的相应部分,这里不再赘述。

[0237] 所述处理器1003用于执行上述申请实施例中由第一终端设备进行的处理。例如,处理器1003用于基于接收到的第一标记报文或第二标记报文执行相应地拥塞响应。具体过程可参见上述本申请实施例相应地部分,这里不再赘述。

[0238] 存储器1001包括操作系统和应用程序,用于存储报文处理的程序、代码或指令,当处理器1003或硬件设备在执行报文处理时,调用并执行这些程序、代码或指令时可以完成图3-图7中涉及第一终端设备的处理过程。具体过程可参见上述本申请实施例相应地部分,这里不再赘述。

[0239] 可以理解的是,图10仅仅示出了第一终端设备的简化设计。在实际应用中,第一终端设备可以包含任意数量的接口,处理器,存储器等,而所有可以实现本申请的第一终端设备都在本申请的保护范围之内。

[0240] 另外,本申请实施例提供了一种计算机存储介质,用于储存为上述第一终端设备所用的计算机软件指令,其包含用于执行上述图3-图7所示实施例所设计的程序。

[0241] 如图11所示,为本申请实施例公开的用于实现报文处理方法的网络系统1100,主要包括:第一终端设备1101和网络设备1102。

[0242] 以上本申请实施例公开的网络系统中,网络设备1102可以具体为图7和图9中公开的网络设备,具体过程以及执行原理可以参照上述说明,这里不再进行赘述。

[0243] 以上本申请实施例公开的网络系统中,第一终端设备1101可以具体为图8和图10中公开的第一终端设备,具体过程以及执行原理可以参照上述说明,这里不再进行赘述。

[0244] 综上所述,本申请实施例公开的一种报文处理的技术方案,网络设备在检测到发

生拥塞后,将获得来自第一终端设备的第一IP报文的源IP地址和目的IP地址交换后,存储于第一数据表项中的源IP地址和目的IP地址中,基于此,获得与第一数据表项中的源IP地址和目的IP地址相同的第二IP报文,并设置CE标记,该第二IP报文为第二终端设备发送的。通过在反向传输的IP报文中设置CE标记,在拥塞探测时,不需要发送给对端设备进行响应,也不需要依赖终端设备自身判断对应该报文的计时是否超时,减少了拥塞探测的时间,进而降低时延。

[0245] 本领域技术人员应该可以意识到,在上述一个或多个示例中,本申请所描述的功能可以用硬件、软件、固件或它们的任意组合来实现。当使用软件实现时,可以将这些功能存储在计算机可读介质中或者作为计算机可读介质中的一个或多个指令或代码进行传输。计算机可读介质包括计算机存储介质和通信介质,其中,通信介质包括便于从一个地方向另一个地方传送计算机程序的任何介质。存储介质可以是通用或专用计算机能够存取的任何可用介质。

[0246] 本说明书的各个部分均采用递进的方式进行描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点介绍的都是与其他实施例不同之处。尤其,对于装置和系统实施例而言,由于其基本相似于方法实施例,所以描述的比较简单,相关之处参见方法实施例部分的说明即可。

[0247] 最后应说明的是:以上实施例仅用以示例性说明本申请的技术方案,而非对其限制;尽管参照前述实施例对本申请及本申请带来的有益效果进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本申请权利要求的范围。



图1

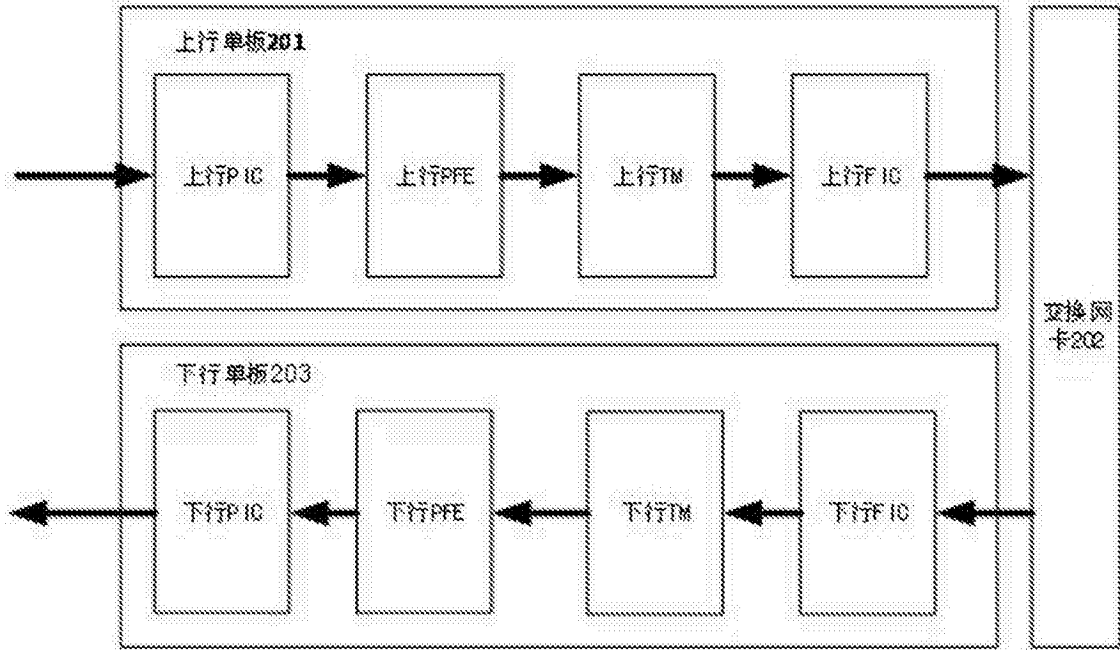


图2

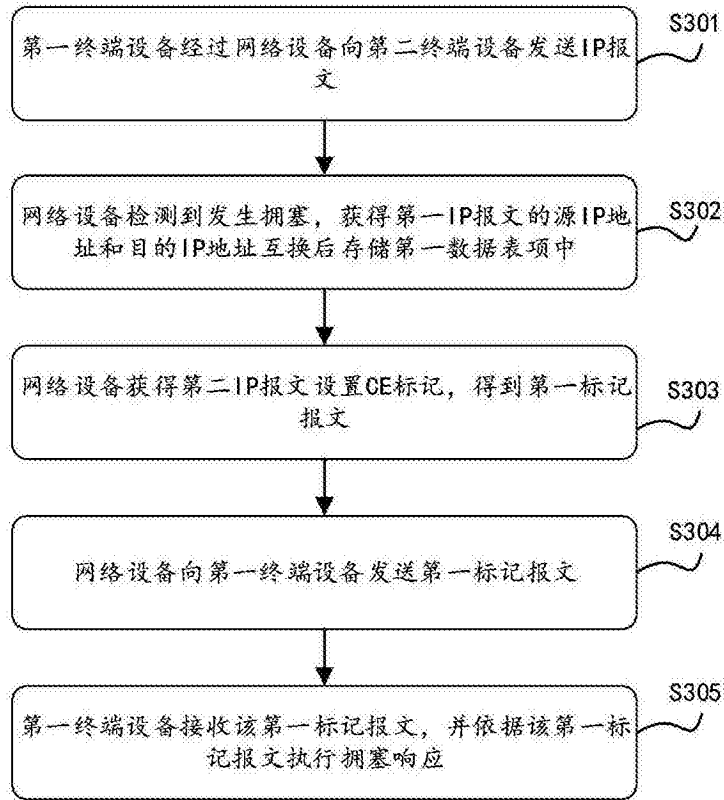


图3

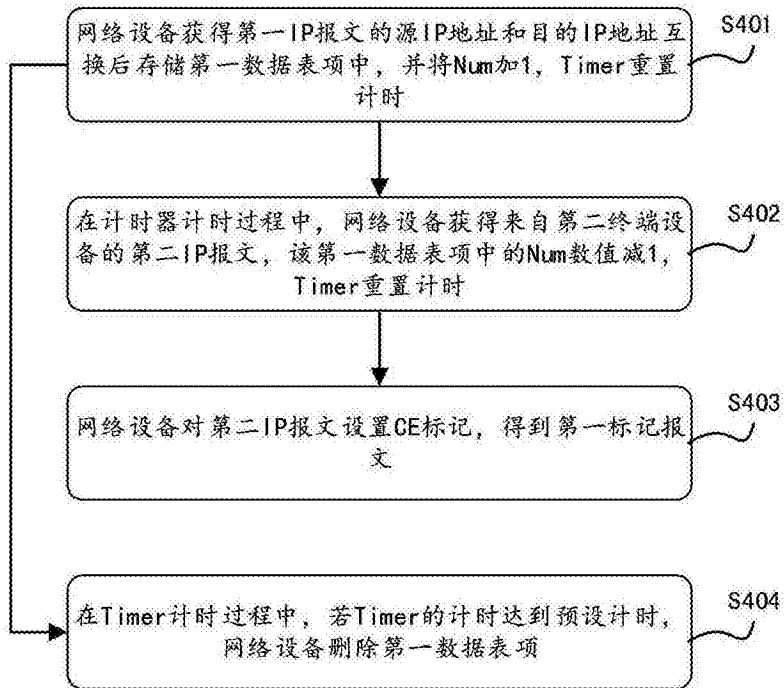


图4

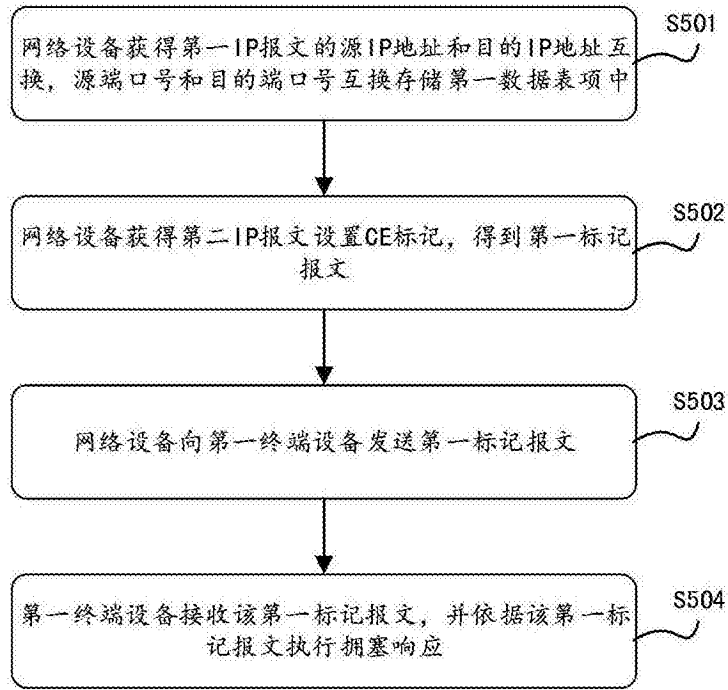


图5

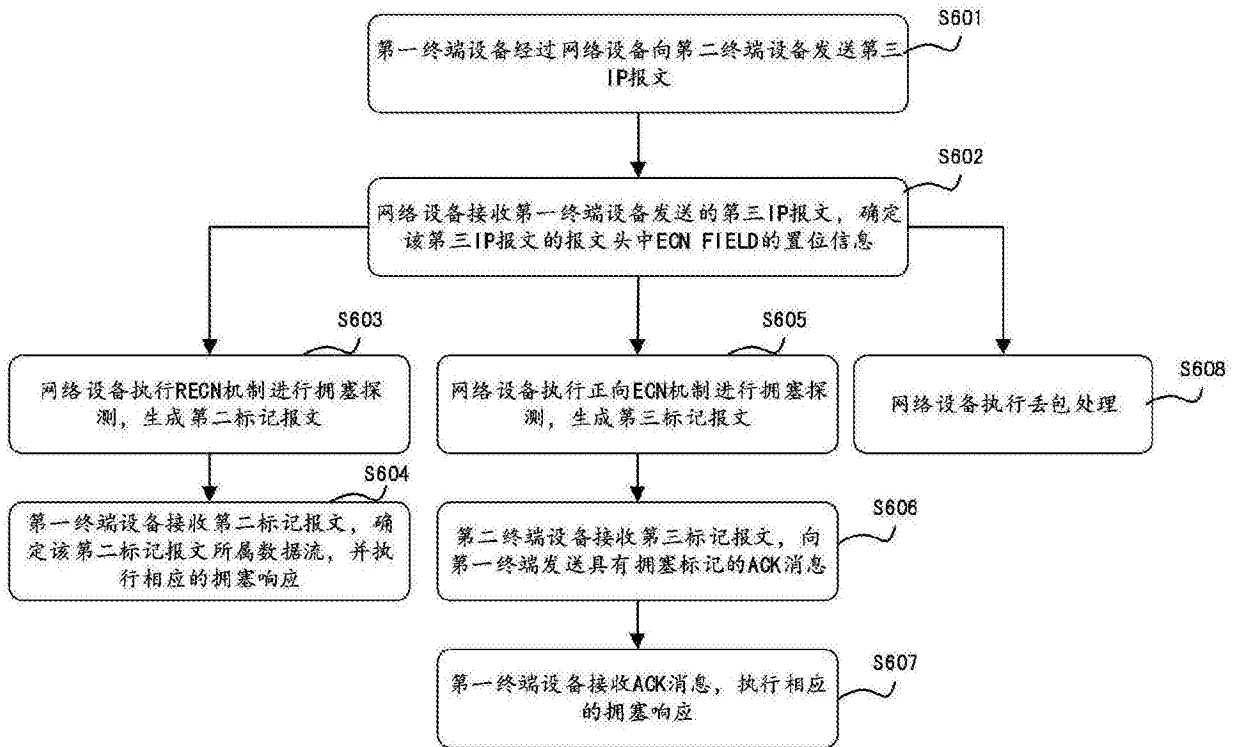


图6

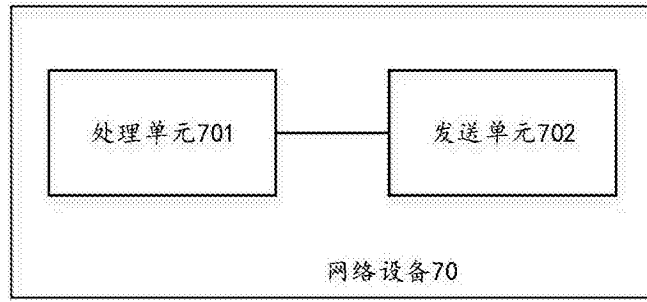


图7

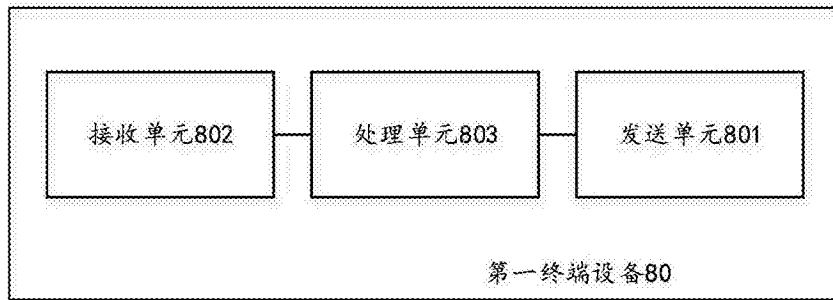


图8

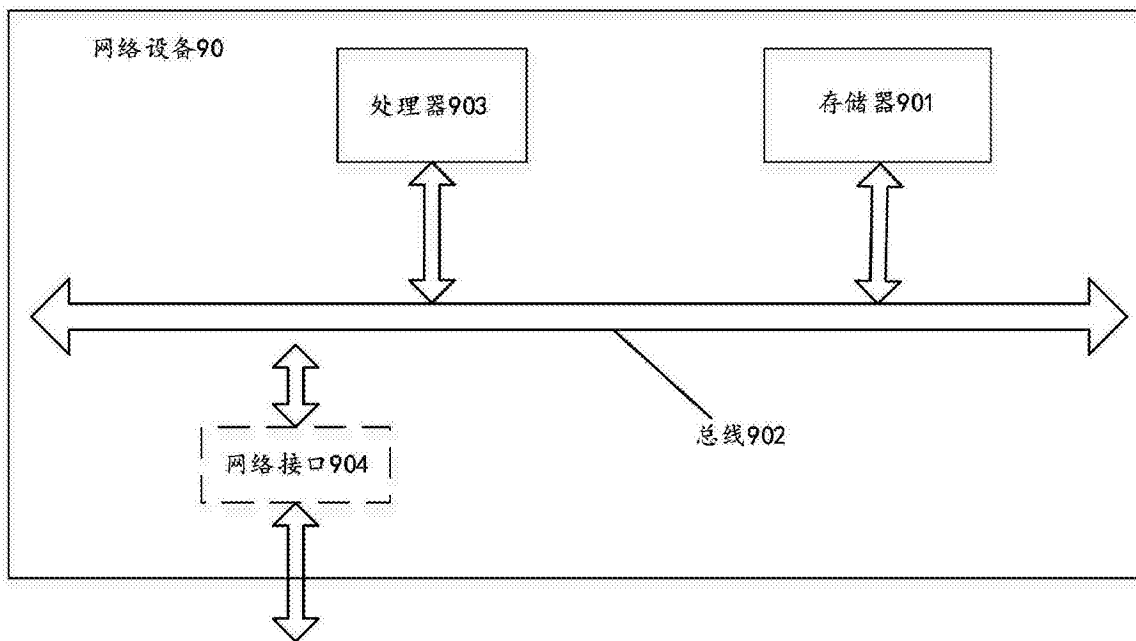


图9

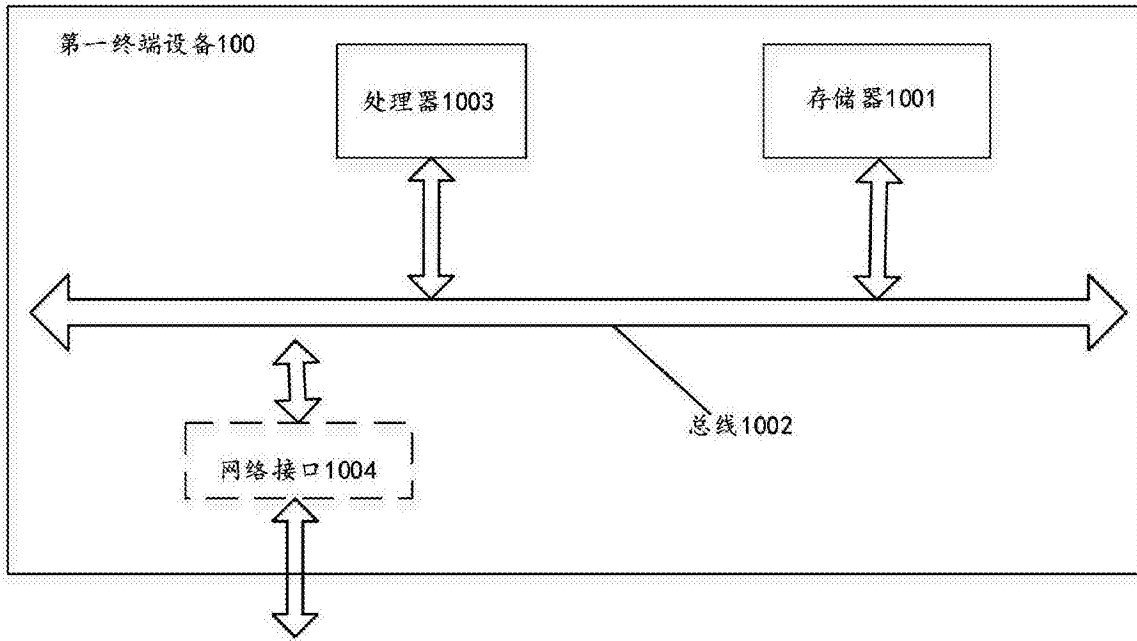


图10

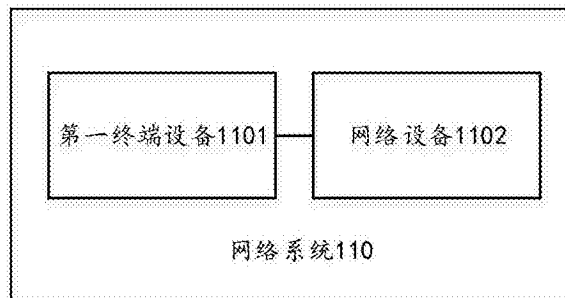


图11