(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 111654447 B
(45)授权公告日 2023.04.18

(21)申请号 202010305935.6

(22)申请日 2018.01.16

(65)同一申请的已公布的文献号
申请公布号 CN 111654447 A

(43)申请公布日 2020.09.11

(62)分案原申请数据
201880003454.0 2018.01.16

(73)专利权人 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 苏德现 杨华

(74)专利代理机构 广州三环专利商标代理有限公司 44202
专利代理师 易浩球

(51)Int.Cl.
H04L 47/125(2022.01)
H04L 45/74(2022.01)

(56)对比文件
US 2010082766 A1,2010.04.01
WO 2005099375 A2,2005.10.27
CN 107231316 A,2017.10.03
US 2009077567 A1,2009.03.19
CN 101409715 A,2009.04.15
CN 101702689 A,2010.05.05

审查员 王刚

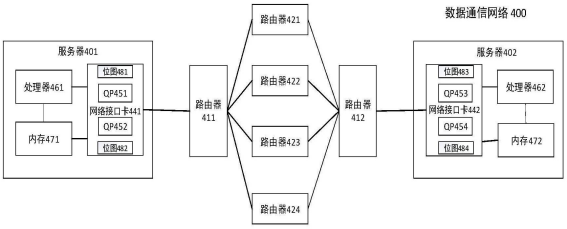
权利要求书4页 说明书17页 附图11页

(54)发明名称

一种报文传输的方法及装置

(57)摘要

本发明提出了一种报文传输方法以及实现该方法的装置。在当前的做法中，路由设备根据所转发报文的五元组信息的哈希值选择转发路径。由于不同的队列对所发送的报文的五元组信息，经过哈希计算的值可能相同或者对应同一条转发路径，导致报文在各个路径上的流量不均衡。针对目前做法所可能导致的流量不均衡的问题，本发明将待发送的报文分为若干分组，不同分组中的报文具有不同的源端口信息，且每个报文携带的报头携带有该报文在目的服务器内的内存中的写入地址。通过这种做法，使得待发送报文会被分在不同的路径上进行转发，从而增加了网络流量的均衡性。



CN 111654447 B

1.一种报文传输的方法,其特征在于,所述方法应用于源端设备,所述源端设备与目的端设备之间通过以太网进行远程直接内存访问RDMA,所述源端设备包括网络接口卡;所述方法包括:

获取Q个数据段;

分别封装所述Q个数据段得到Q个报文,并分别发送所述Q个报文,其中,所述Q个报文中的每个报文携带第一报头和第二报头,每个报文携带的第一报头用于指示本报文在所述目的端设备的内存中的写入地址,每个报文携带的第二报头包含源端口号信息,所述Q个报文中至少两个报文分别携带的第二报头中的源端口号信息不相同;

所述Q为大于等于2的正整数。

2.根据权利要求1所述的方法,其特征在于,所述分别封装所述Q个数据段得到Q个报文,并发送所述Q个报文包括:

根据源端口号信息,依次封装所述Q个数据段得到所述Q个报文,每封装完成一个报文就发送封装后的报文,并在每封装完成N个报文后,更新源端口号信息,前一组N个报文携带的源端口号信息与后一组N个报文携带的源端口号信息不同,N大于等于1,小于Q。

3.根据权利要求1所述的方法,其特征在于,所述分别封装所述Q个数据段得到Q个报文,并发送所述Q个报文包括:

将所述Q个数据段划分为M个分组,每个分组中包括至少一个数据段,依次封装每个分组中的数据段得到每个分组中的报文,其中,每个分组中的报文携带的源端口号信息相同,至少两个分组中的报文携带的源端口号信息不同,M小于等于Q。

4.根据权利要求1-3任意一项所述的方法,其特征在于,所述分别封装所述Q个数据段得到Q个报文之前,还包括:

根据所述Q个数据段的第一个数据段的基地址和每个数据段的长度,确定所述Q个报文中的每个报文在所述目的端设备的内存中的写入地址。

5.根据权利要求1-4任意一项所述的方法,其特征在于,所述Q个报文中的每个报文还分别携带报文序号,每个报文携带的报文序号用于指示本报文在所述Q个报文中的发送顺序。

6.一种报文传输的方法,其特征在于,所述方法应用于目的端设备,源端设备与所述目的端设备之间通过以太网进行远程直接内存访问RDMA,所述目的端设备包括网络接口卡;所述方法包括:

接收Q个报文,其中,每个报文携带第一报头和第二报头,每个报文携带的第一报头用于指示本报文在所述目的端设备的内存中的写入地址,每个报文携带的第二报头包含源端口号信息,所述Q个报文中至少两个报文携带的第二报头中的源端口号信息不相同,Q为大于等于2的正整数;

根据所述Q个报文各自携带的本报文在所述目的端设备的内存中的写入地址,分别将所述Q个报文保存到所述目的端设备的内存中。

7.根据权利要求6所述的方法,其特征在于,所述接收Q个报文包括:依次接收Q个报文;则,所述保存所述Q个报文到所述目的端设备的内存中包括:每接收到一个报文,就执行将接收到的报文保存到所述目的端设备的内存中的步骤。

8.根据权利要求7所述的方法,其特征在于,所述Q个报文中的每个报文还分别携带报

文序号，每个报文携带的报文序号用于指明本报文在所述Q个报文中的发送顺序；

所述方法还包括：

每接收到所述Q个报文中的一个报文，记录当前接收到的报文携带的报文序号，并根据当前接收到的报文的报文序号，确定预备接收的下一个报文的报文序号；

在接收到所述Q个报文中的下一个报文后，确定接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致，如果否，启动丢包检测流程；

如果通过所述丢包检测流程确定在报文传输过程中发生丢包，则向所述源端设备发送报文重传指示。

9.根据权利要求8所述的方法，其特征在于，所述网络接口卡设置有位图，所述位图包括至少Q个位图位，所述Q个位图位按照所述Q个报文的发送顺序从前往后对应于所述Q个报文，所述位图设置有头指针和尾指针，所述头指针指向所述网络接口卡最新接收到的报文所对应的位图位，所述尾指针指向所述网络接口卡预备接收的下一个报文；

所述每接收到一个报文，记录当前接收到的报文携带的报文序号，并根据当前接收到的报文的报文序号，确定预备接收的下一个报文的报文序号，包括：

根据当前接收到的报文的报文序号，将所述位图中代表所述当前接收到的报文的位图位设置为有效，并将所述头指针指向代表所述当前接收到的报文的位图位；以及，

根据所述当前接收到的报文的报文序号，确定所述当前接收到的报文是否是所述尾指针当前指向的位图位所对应的报文，如果是，更新所述尾指针的指向，所述尾指针新的指向为所述当前接收到的报文所对应的位图位之后的无效的位图位中的第一个位图位，如果否，保持所述尾指针当前指向的位图位不变。

10.根据权利要求9所述的方法，其特征在于，所述确定接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致，包括：

根据所述接收到的下一个报文的报文序号，确定所述尾指针当前是否指向所述接收到的下一个报文所对应的位图位。

11.根据权利要求9或10所述的方法，其特征在于，所述丢包检测流程包括：

针对所述尾指针当前指向的位图位所对应的报文启动定时器，若在定时器超时后，所述尾指针的指向不发生改变，确定所述尾指针当前指向的位图位所对应的报文发生丢包。

12.根据权利要求9或10所述的方法，其特征在于，所述丢包检测流程包括：

确定所述头指针当前指向的位图位是否超过预定值，如果超过，确定所述头指针和所述尾指针之间的位图位所对应的报文发生丢包。

13.根据权利要求9至12任一项所述的方法，其特征在于，所述向所述源端设备发送报文重传指示包括：

向所述源端设备发送报文重传指示，所述重传指示携带所述尾指针当前指向的位图位所对应的报文的报文序号，以请求所述源端设备将所述Q个报文中所述尾指针当前指向的位图位所对应的报文之后的所有报文进行重新发送。

14.一种网络接口卡，其特征在于，所述网络接口卡位于远程直接内存访问RDMA的源端设备；

所述网络接口卡包括：

获取模块，用于获取Q个数据段；

发送模块，用于分别封装所述Q个数据段得到Q个报文，并分别发送所述Q个报文，其中，所述Q个报文中的每个报文携带第一报头和第二报头，每个报文携带的第一报头用于指示本报文在目的端设备的内存中的写入地址，每个报文携带的第二报头包含源端口号信息，所述Q个报文中至少两个报文分别携带的第二报头中的源端口号信息不相同，Q为大于等于2的正整数。

15.根据权利要求14所述的网络接口卡，其特征在于，所述发送模块具体用于根据源端口号信息，依次封装所述Q个数据段得到所述Q个报文，每封装完成一个报文就发送封装后的报文，并在每封装完成N个报文后，更新源端口号信息，前一组N个报文携带的源端口号信息与后一组N个报文携带的源端口号信息不同，N大于等于1，小于Q。

16.根据权利要求14所述的网络接口卡，其特征在于，所述发送模块具体用于将所述Q个数据段划分为M个分组，每个分组中包括至少一个数据段，依次封装每个分组中的数据段得到每个分组中的报文，其中，每个分组中的报文携带的源端口号信息相同，至少两个分组中的报文携带的源端口号信息不同，M小于等于Q。

17.根据权利要求14-16任意一项所述的网络接口卡，其特征在于，还包括：

确定模块，用于根据所述Q个数据段的第一个数据段的基地址和每个数据段的长度，确定所述Q个报文中的每个报文在所述目的端设备的内存中的写入地址。

18.根据权利要求14-17任意一项所述的网络接口卡，其特征在于，所述Q个报文中的每个报文还分别携带报文序号，每个报文携带的报文序号用于指示本报文在所述Q个报文中的发送顺序。

19.一种计算设备，其特征在于，所述设备包括主处理系统和网络接口卡；

所述主处理系统用于处理业务，在需要将业务数据发送到目的端设备时，将所述业务数据发送到所述网络接口卡中；

所述网络接口卡，用于获取Q个数据段，所述Q个数据段属于所述业务数据，分别封装所述Q个数据段得到Q个报文，并分别发送所述Q个报文，其中，所述Q个报文中的每个报文携带第一报头和第二报头，每个报文携带的第一报头用于指示本报文在所述目的端设备的内存中的写入地址，每个报文携带的第二报头包含源端口号信息，所述Q个报文中至少两个报文分别携带的第二报头中的源端口号信息不相同，Q为大于等于2的正整数。

20.根据权利要求19所述的设备，其特征在于，所述网络接口卡还用于根据所述Q个数据段的第一个数据段的基地址和每个数据段的长度，确定所述Q个报文中的每个报文在所述目的端设备的内存中的写入地址。

21.一种网络接口卡，其特征在于，所述网络接口卡位于远程直接内存访问RDMA的目的端设备；

所述网络接口卡包括：

接收模块，用于接收Q个报文，其中，每个报文携带第一报头和第二报头，每个报文携带的第一报头用于指示本报文在目的端设备的内存中的写入地址，每个报文携带的第二报头包含源端口号信息，所述Q个报文中至少两个报文分别携带的第二报头中的源端口号信息不相同，Q为大于等于2的正整数；

执行模块，用于根据所述Q个报文各自携带的本报文在所述目的端设备的内存中的写入地址，分别将所述Q个报文保存到所述目的端设备的内存中。

22.根据权利要求21所述的网络接口卡，其特征在于，所述接收模块具体用于：依次接收Q个报文，所述接收模块每接收到一个报文，所述执行模块就执行将接收到的报文保存到所述目的端设备的内存中的步骤。

23.根据权利要求22所述的网络接口卡，其特征在于，所述Q个报文中的每个报文还分别携带报文序号，每个报文携带的报文序号用于指明本报文在所述Q个报文中的发送顺序；

所述网络接口卡还包括：检测模块；所述接收模块每接收到一个报文，所述检测模块用于记录当前接收到的报文携带的报文序号，并根据当前接收到的报文的报文序号，确定预备接收的下一个报文的报文序号；在接收到下一个报文后，确定接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致，如果否，启动丢包检测流程；如果通过所述丢包检测流程确定在报文传输过程中发生丢包，则向源端设备发送报文重传指示。

24.根据权利要求23所述的网络接口卡，其特征在于，所述网络接口卡设置有位图，所述位图至少包括Q个位图位，所述Q个位图位按照所述Q个报文的发送顺序对应于所述Q个报文，所述位图设置有头指针和尾指针，所述头指针指向所述网络接口卡最新接收到的报文所对应的位图位，所述尾指针指向所述网络接口卡预备接收的下一个报文；

所述检测模块具体用于根据当前接收到的报文的报文序号，将所述位图中代表所述当前接收到的报文的位图位设置为有效，并将所述头指针指向代表所述当前接收到的报文的位图位；以及，根据所述当前接收到的报文的报文序号，确定所述当前接收到的报文是否是所述尾指针当前指向的位图位所对应的报文，如果是，更新所述尾指针的指向，所述尾指针新的指向为所述当前接收到的报文所对应的位图位之后的无效的位图位中的第一个位图位，如果否，保持所述尾指针当前指向的位图位不变。

25.根据权利要求24所述的网络接口卡，其特征在于，所述检测模块执行丢包检测流程具体包括：针对所述尾指针当前指向的位图位所对应的报文启动定时器，若在定时器超时后，所述尾指针的指向不发生改变，确定所述尾指针当前指向的位图位所对应的报文发生丢包。

26.一种设备，其特征在于，所述设备包括主处理系统和网络接口卡；

所述主处理系统用于从所述设备的内存中获取应用数据，以及根据所述应用数据处理业务；

所述网络接口卡，用于执行如权利要求6-13任意一项所述的方法。

27.一种计算机存储介质，其特征在于，包括指令，当其在计算机上运行时，使得计算机执行如权利要求1-5任意一项所述的方法。

28.一种计算机存储介质，其特征在于，包括指令，当其在计算机上运行时，使得计算机执行如权利要求6-13任意一项所述的方法。

一种报文传输的方法及装置

技术领域

[0001] 本发明涉及报文传输技术领域,特别涉及一种报文传输的方法及装置。

背景技术

[0002] 在数据通信系统,为了提高服务器之间报文传输的速度,通常采用远程直接内存访问(英文:Remote Direct Memory Access,简称:RDMA)技术进行连接。RDMA,是通过网络把数据直接传入计算机的存储区,将数据从一个系统快速移动到远程系统存储器中,而不对操作系统造成影响。RDMA消除了外部存储器复制和上下文切换的开销,因此能解放内存带宽和CPU周期用于改进应用系统性能。

[0003] 基于融合以太网的远程直接内存访问(英文全称:RDMA over Converged Ethernet,缩写:RoCE)是RDMA技术的一种,允许服务器通过以太网进行远程直接内存访问。目前RoCE有两个协议版本,v1和v2。其中,RoCE v1协议允许在同一个广播域下的任意两台服务器直接访问。而RoCE v2协议则可以实现路由功能。虽然RoCE协议的优点主要是基于融合以太网的特性,但是RoCE协议也可以应用在传统以太网网络或者非融合以太网络中。

[0004] 当RoCEv2协议的报文在多路径的网络中进行传输时,通常根据该报文中的五元组信息的哈希值来选择转发的路径,以此实现流量均衡。然而,根据RoCEv2协议的快启动特性,从某个源端端口发出的报文流量有可能在某个时间段可能比较大,另外,哈希的随机性也可能造成多路径网络中的某条路径在某时刻的流量较大,这都可能导致多路径网络中发生某条路径的拥塞。当网络产生拥塞之后,不仅会导致网络时延增加,也会导致网络丢包的可能性增加,从而导致网络传输的有效带宽下降。RoCE协议下的网络路由的路径均衡需要进一步优化。

发明内容

[0005] 本申请的实施例提供一种报文传输方法,以使得采用RoCE协议的报文在以太网中实现更均衡的路由传输。

[0006] 第一方面,本申请提供了一种报文传输的方法,该方法应用于数据通信系统,该数据通信系统中的源端设备与目的端设备之间通过以太网进行远程直接内存访问RDMA,源端设备的网络接口卡上包括至少源队列对,源队列对包括发送队列。该报文传输的方法包括:从源队列对的发送队列中获取Q个数据段;分别封装Q个数据段得到Q个报文,并分别发送该Q个报文,其中,这Q个报文中的每个报文携带第一报头和第二报头,每个报文携带的第一报头用于指示该报文在目的端设备的内存中的写入地址,每个报文携带的第二报头包含源端口号信息,Q个报文中至少两个报文分别携带的第二报头中的源端口号信息不相同,Q为大于等于2的正整数。

[0007] 上述方案中,由于至少两个报文携带的第二报头中的源端口号信息不同,因此路由器根据五元组的哈希值进行选路时,该组报文分在至少两个不同的网络路径中进行传输,从而使得网络中各个路径的流量较为均衡。另一方面,由于相同的一组报文在不同传输

路径传输可能导致目的端接收到乱序的一组报文，上述方案通过在报文中携带具有指示该报文在目的端的内存中的写入地址的第一报头，可以使得目的端设备能够直接根据每个报文携带的地址信息进行RDMA操作，从而使得上述方案既能够实现RDMA操作的报文的路由的进一步优化，又能够保证RDMA操作能够在目的端真正实现。

[0008] 对于上述第一方面，一种可能的实现方式是根据源队列对配置的源端口号信息，依次封装Q个数据段得到Q个报文，每封装完成一个报文就发送封装后的报文，并在每封装完成N个报文后，更新源队列对配置的源端口号信息，前一组N个报文携带的源端口号信息与后一组N个报文携带的源端口号信息不同，N大于等于1，小于Q。通过上述每封装一个报文就发送该报文的做法，可以提高系统的效率。

[0009] 对于上述第一方面，另一种可能的实现方式是：将Q个数据段划分为M个分组，每个分组中包括至少一个数据段，依次封装每个分组中的数据段得到每个分组中的报文，其中，每个分组中的报文携带的源端口号信息相同，至少两个分组中的报文携带的源端口号信息不同。通过上述分组封装的方法，可以提高系统的效率。

[0010] 对于上述第一方面，另一种可能的实现方式是分别封装Q个数据段得到Q个报文之前，还包括：根据Q个数据段的第一个数据段的基地址和每个数据段的长度，确定Q个报文中的每个报文在目的端设备的内存中的写入地址。采取计算每个报文在目的端设备的内存中的写入地址并将该地址封装在报文中的做法，可以使得报文在到达目的端时直接被写入内存相应的地址中。

[0011] 对于上述第一方面，另一种可能的实现方式是Q个报文中的每个报文还分别携带报文序号，每个报文携带的报文序号用于指示本报文在Q个报文中的发送顺序。通过这种做法，可以便于目的端根据报文序号确认该组报文是否收齐或者进行报文的乱序重排等，提高了系统的稳定性。

[0012] 第二方面，提供一种报文传输的方法，该方法应用于数据通信系统，该数据通信系统中的源端设备和目的端设备之间通过以太网进行远程直接内存访问RDMA，其中，目的端设备的网络接口卡上包括目的队列对，目的队列对包括接收队列；该报文传输的方法包括：接收Q个报文，其中，每个报文携带第一报头和第二报头，每个报文携带的第一报头用于指示本报文在目的端设备的内存中的写入地址，每个报文携带的第二报头包含源端口号信息，Q个报文中至少两个报文携带的第二报头中的源端口号信息不相同，Q为大于等于2的正整数；根据Q个报文各自携带的本报文在目的端设备的内存中的写入地址，分别将Q个报文从目的队列对保存到目的端设备的内存中。

[0013] 对于源端发送的一组报文，由于其在多路径网络中路由可能经过不同的传输路径，因此达到目的端的顺序与源端的发送顺序可能不同，目的端在接收到源端发送的报文后，直接根据报文携带的写入地址进行内存写入，而不是等待接收到全部一组报文后进行重排之后才进行内存写入，提高了系统效率，同时也避免了若一组报文在传输中发生丢包，将可能全部一组的报文将无法实现目的端内存写入的问题。

[0014] 对于上述第二方面，一种可能的实现方式是，接收Q个报文包括：依次接收Q个报文；保存Q个报文到目的端设备的内存中包括：每接收到一个报文，就执行将接收到的报文保存到目的端设备的内存中的步骤。通过这种做法，可以每接收一个报文就进行相应的处理，提高了系统的效率。

[0015] 对于上述第二方面，另一种可能的实现方式是，Q个报文中的每个报文还分别携带报文序号，每个报文携带的报文序号用于指明本报文在Q个报文中的发送顺序。该实现方式还包括：每接收到一个报文，记录当前接收到的报文携带的报文序号，并根据当前接收到的报文的报文序号，确定预备接收的下一个报文的报文序号；在接收到下一个报文后，确定接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致，如果否，启动丢包检测流程；如果通过丢包检测流程确定在报文传输过程中发生丢包，则向源端设备发送报文重传指示。通过这种做法，可以当乱序、丢包等情况发生时，避免立刻向源端发送重传指示，而是启动相应的丢包检测，在丢包检测确定发生丢包的情况下，才引导源端进行报文重传，提升了系统的稳定性。

[0016] 对于上述第二方面，另一种可能的实现方式是目的队列对设置有位图，该位图至少包括Q个位图位，该Q个位图位按照Q个报文的发送顺序从前往后对应于该Q个报文，位图设置有头指针和尾指针，头指针指向接收队列最新接收到的报文所对应的位图位，尾指针指向接收队列预备接收的下一个报文；每接收到一个报文，记录当前接收到的报文携带的报文序号，并根据当前接收到的报文的报文序号，确定预备接收的下一个报文的报文序号，包括：根据当前接收到的报文的报文序号，将位图中代表当前接收到的报文的位图位设置为有效，并将头指针指向代表当前接收到的报文的位图位；以及，根据当前接收到的报文的报文序号，确定当前接收到的报文是否是尾指针当前指向的位图位所对应的报文，如果是，更新尾指针的指向，尾指针新的指向为所述当前接收到的报文所对应的位图位之后的无效的位图位中的第一个位图位，如果否，保持尾指针当前指向的位图位不变。通过这种做法，利用位图来统计接收到的报文状况，提高了系统的效率。

[0017] 对于上述第二方面，另一种可能的实现方式是，确认接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致，包括：根据接收到的下一个报文的报文序号，确定尾指针当前是否指向接收到的下一个报文所对应的位图位。通过这种做法，可以判断接收到的报文是否发生乱序，从而决定是否要采取相应的措施。

[0018] 对于上述第二方面，另一种可能的实现方式是，丢包检测流程包括：针对尾指针当前指向的位图位所对应的报文启动定时器，若在定时器超时后，尾指针的指向不发生改变，确定尾指针当前指向的位图位所对应的报文发生丢包。通过这种做法，当某个报文一直没有被接收到时，系统可以判定该报文丢失，提高了系统的效率。

[0019] 对于上述第二方面，另一种可能的实现方式是，丢包检测流程包括：确定头指针当前指向的位图位是否超过预定值，如果超过，确定头指针和尾指针之间的位图位所对应的报文发生丢包。通过这种做法，可以有效判断接收到的报文是否有丢包产生。

[0020] 对于上述第二方面，另一种可能的实现方式是，向源端设备发送报文重传指示包括：向源端设备发送报文重传指示，该重传指示携带尾指针当前指向的位图位所对应的报文的报文序号，以请求源端设备将Q个报文中的尾指针当前指向的位图位所对应的报文之后的所有报文进行重新发送。通过这种做法，只需要源端重传尾指针当前指向的位图位所对应的报文之后的所有报文，提高了系统的效率。

[0021] 对于上述第二方面，另一种可能的实现方式是当一组报文所对应的位图位的值都被置为有效时，说明该组报文已经全部收齐，目的端向源端发送确认应答报文。通过这种做法，可以判断何时一组报文已经全部收齐。

[0022] 对于上述第二方面，另一种可能的实现方式是当目的端接收到的报文没有携带含有指示该报文在目的端的写入地址的部分时，先缓存该报文，并确认报文是否发生乱序、丢包以及是否收齐。当确认整组报文都已经收齐后，根据报文的报文序号进行乱序重排，并在乱序重排后，将报文写入内存中。通过这种做法，可以接收没有携带含有指示该报文在目的端的写入地址的报文，并可以进行乱序重排。

[0023] 第三方面，提供一种网络接口卡，该网络接口卡位于远程直接内存访问RDMA的源端设备，该网络接口卡上设置有源队列对，每队列对包括发送队列；该网络接口卡包括：获取模块，用于从源队列对的发送队列中获取Q个数据段；发送模块，用于封装Q个数据段得到Q个报文，并发送Q个报文，其中，Q个报文中的每个报文携带第一报头、第二报头和队列对标识，每个报文携带的第一报头用于指示本报文在目的端设备的内存中的写入地址，每个报文携带的第二报头包含源端口号信息，Q个报文中至少两个报文携带的第二报头中的源端口号信息不相同，Q为大于等于2的正整数。

[0024] 对于上述第三方面，一种可能的实现方式是，发送模块具体用于根据源队列对配置的源端口号信息，依次封装Q个数据段得到Q个报文，每封装完成一个报文就发送封装后的报文，并在每封装完成N个报文后，更新源队列对配置的源端口号信息，前一组N个报文携带的源端口号信息与后一组N个报文携带的源端口号信息不同，N大于等于1，小于Q。

[0025] 对于上述第三方面，另一种可能的实现方式是，发送模块具体用于将Q个数据段划分为M个分组，每个分组中包括至少一个数据段，依次封装每个分组中的数据段得到每个分组中的报文，其中，每个分组中的报文携带的源端口号信息相同，至少两个分组中的报文携带的源端口号信息不同，M小于等于Q。

[0026] 对于上述第三方面，另一种可能的实现方式是，还包括确定模块，用于根据Q个数据段的第一个数据段的基地址和每个数据段的长度，确定Q个报文中的每个报文在目的端设备的内存中的写入地址。

[0027] 对于上述第三方面，另一种可能的实现方式是，Q个报文中的每个报文还分别携带报文序号，每个报文携带的报文序号用于指示本报文在Q个报文中的发送顺序。

[0028] 第四方面，提供一种设备，该设备包括主处理系统和网络接口卡；主处理系统用于处理业务，在需要将业务数据发送到目的端设备时，将业务数据发送到网络接口卡中的业务数据对应的源队列对的发送队列；网络接口卡用于从业务数据对应的源队列对的发送队列中获取Q个数据段，该Q个数据段属于业务数据，封装该Q个数据段得到Q个报文，并发送Q个报文，其中，Q个报文中的每个报文携带第一报头和第二报头，每个报文携带的第一报头用于指示本报文在目的端设备的内存中的写入地址，每个报文携带的第二报头包含源端口号信息，Q个报文中至少两个报文携带的第二报头中的源端口号信息不相同，Q为大于等于2的正整数。

[0029] 对于上述第四方面，一种可能的实现方式是，网络接口卡封装Q个数据段得到Q个报文，并发送Q个报文包括：根据源队列对配置的源端口号信息，依次封装Q个数据段得到Q个报文，每封装完成一个报文就发送封装后的报文，并在每封装完成N个报文后，更新源队列对配置的源端口号信息，前一组N个报文携带的源端口号信息与后一组N个报文携带的源端口号信息不同，N大于等于1，小于Q。

[0030] 对于上述第四方面，另一种可能的实现方式是，网络接口卡封装Q个数据段得到Q

个报文，并发送Q个报文包括：将Q个数据段划分为M个分组，每个分组中包括至少一个数据段，依次封装每个分组中的数据段得到每个分组中的报文，其中，每个分组中的报文携带的源端口号信息相同，至少两个分组中的报文携带的源端口号信息不同。

[0031] 对于上述第四方面，另一种可能的实现方式是，网络接口卡还用于根据Q个数据段的第一个数据段的基地址和每个数据段的长度，确定Q个报文中的每个报文在目的端设备的内存中的写入地址。

[0032] 对于上述第四方面，另一种可能的实现方式是，网络接口卡在封装Q个报文时，Q个报文中的每个报文还分别携带报文序号，每个报文携带的报文序号用于指示本报文在Q个报文中的发送顺序。

[0033] 第五方面，提供一种网络接口卡，该网络接口卡位于远程直接内存访问RDMA的目的端设备，该网络接口卡上设置有目的队列对，目的队列对包括接收队列；该网络接口卡包括：接收模块，用于接收Q个报文，其中，每个报文携带第一报头和第二报头，每个报文携带的第一报头用于指示本报文在目的端设备的内存中的写入地址，每个报文携带的第二报头包含源端口号信息，Q个报文中至少两个报文携带的第二报头中的源端口号信息不相同，Q为大于等于2的正整数，目的端设备为RDMA的目的端设备；执行模块，用于根据Q个报文各自携带的本报文在目的端设备的内存中的写入地址，分别将Q个报文从目的队列对保存到目的端设备的内存中。

[0034] 对于上述第五方面，一种可能的实现方式是接收模块具体用于依次接收Q个报文；在接收模块每接收到一个报文，执行模块就执行将接收到的报文保存到目的端设备的内存中的步骤。

[0035] 对于上述第五方面，另一种可能的实现方式是，Q个报文中的每个报文还分别携带报文序号，每个报文携带的报文序号用于指明本报文在所述Q个报文中的发送顺序；网络接口卡还包括：检测模块；在接收模块每接收到一个报文，检测模块用于记录当前接收到的报文携带的报文序号，并根据当前接收到的报文的报文序号，确定预备接收的下一个报文的报文序号；在接收到下一个报文后，确定接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致，如果否，启动丢包检测流程；如果通过丢包检测流程确定在报文传输过程中发生丢包，则向源端设备发送报文重传指示。

[0036] 对于上述第五方面，另一种可能的实现方式是，目的队列对设置有位图，位图至少包括Q个位图位，该Q个位图位按照Q个报文的发送顺序对应于Q个报文，位图设置有头指针和尾指针，头指针指向本队列的接收队列最新接收到的报文所对应的位图位，尾指针指向本队列对的接收队列预备接收的下一个报文；检测模块具体用于根据当前接收到的报文的报文序号，将位图中代表当前接收到的报文的位图位设置为有效，并将头指针指向代表当前接收到的报文的位图位；以及，根据当前接收到的报文的报文序号，确定当前接收到的报文是否是尾指针当前指向的位图位所对应的报文，如果是，更新尾指针的指向，尾指针新的指向为当前接收到的报文所对应的位图位之后的无效的位图位中的第一个位图位，如果否，保持尾指针当前指向的位图位不变。

[0037] 对于上述第五方面，另一种可能的实现方式是，检测模块确认接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致，包括：根据接收到的下一个报文的报文序号，确定尾指针当前是否指向接收到的下一个报文所对应的位图位。对

于上述第五方面，另一种可能的实现方式是，检测模块执行丢包检测流程具体包括：针对尾指针当前指向的位图位所对应的报文启动定时器，若在定时器超时后，尾指针的指向不发生改变，确定尾指针当前指向的位图位所对应的报文发生丢包。

[0038] 对于上述第五方面，另一种可能的实现方式是，检测模块执行丢包检测流程具体包括：确定头指针当前指向的位图位是否超过预定值，如果超过，确定头指针和尾指针之间的位图位所对应的报文发生丢包。

[0039] 对于上述第五方面，另一种可能的实现方式是，检测模块向源端设备发送报文重传指示包括：向源端设备发送报文重传指示，该重传指示携带尾指针当前指向的位图位所对应的报文的报文序号，以请求源端设备将Q个报文中的尾指针当前指向的位图位所对应的报文之后的所有报文进行重新发送。

[0040] 第六方面，提供一种设备，该设备包括主处理系统和网络接口卡，该主处理系统用于从设备的内存中获取应用数据，以及根据该应用数据处理业务；该网络接口卡用于接收Q个报文，其中，每个报文携带第一报头和第二报头，每个报文携带的第一报头用于指示本报文在目的端设备的内存中的写入地址，每个报文携带的第二报头包含源端口号信息，Q个报文中至少两个报文携带的第二报头中的源端口号信息不相同，Q为大于等于2的正整数；根据Q个报文各自携带的本报文在目的端设备的内存中的写入地址，分别将Q个报文从目的队列对保存到目的端设备的内存中。

[0041] 对于上述第六方面，一种可能的实现方式是，网络接口卡接收Q个报文包括：依次接收Q个报文；保存Q个报文到目的端设备的内存中包括：每接收到一个报文，就执行将接收到的报文保存到目的端设备的内存中的步骤。

[0042] 对于上述第六方面，另一种可能的实现方式是，Q个报文中的每个报文还分别携带报文序号，每个报文携带的报文序号用于指明本报文在Q个报文中的发送顺序。该实现方式还包括：每接收到一个报文，记录当前接收到的报文携带的报文序号，并根据当前接收到的报文的报文序号，确定预备接收的下一个报文的报文序号；在接收到下一个报文后，确定接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致，如果否，启动丢包检测流程；如果通过丢包检测流程确定在报文传输过程中发生丢包，则向源端设备发送报文重传指示。

[0043] 对于上述第六方面，另一种可能的实现方式是，目的队列对设置有位图，该位图至少包括Q个位图位，该Q个位图位按照Q个报文的发送顺序从前往后对应于该Q个报文，位图设置有头指针和尾指针，头指针指向接收队列最新接收到的报文所对应的位图位，尾指针指向接收队列预备接收的下一个报文；每接收到一个报文，记录当前接收到的报文携带的报文序号，并根据当前接收到的报文的报文序号，确定预备接收的下一个报文的报文序号，包括：根据当前接收到的报文的报文序号，将位图中代表当前接收到的报文的位图位设置为有效，并将头指针指向代表当前接收到的报文的位图位；以及，根据当前接收到的报文的报文序号，确定当前接收到的报文是否是尾指针当前指向的位图位所对应的报文，如果是，更新尾指针的指向，尾指针新的指向为所述当前接收到的报文所对应的位图位之后的无效的位图位中的第一个位图位，如果否，保持尾指针当前指向的位图位不变。

[0044] 对于上述第六方面，另一种可能的实现方式是，网络接口卡确认接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致，包括：根据接收到的

下一个报文的报文序号，确定尾指针当前是否指向接收到的下一个报文所对应的位图位。

[0045] 对于上述第六方面，另一种可能的实现方式是，网络接口卡进行丢包检测流程包括：针对尾指针当前指向的位图位所对应的报文启动定时器，若在定时器超时后，尾指针的指向不发生改变，确定尾指针当前指向的位图位所对应的报文发生丢包。

[0046] 对于上述第六方面，另一种可能的实现方式是，网络接口卡进行丢包检测流程包括：确定头指针当前指向的位图位是否超过预定值，如果超过，确定头指针和尾指针之间的位图位所对应的报文发生丢包。

[0047] 对于上述第六方面，另一种可能的实现方式是，网络接口卡向源端设备发送报文重传指示包括：向源端设备发送报文重传指示，该重传指示携带尾指针当前指向的位图位所对应的报文的报文序号，以请求源端设备将Q个报文中的尾指针当前指向的位图位所对应的报文之后的所有报文进行重新发送。

[0048] 第七方面，提供一种通信装置，该通信装置包括处理器以及与该处理器耦合的存储器，处理器用于根据存储器中加载的程序指令执行如第一方面所述的报文传输的方法。

[0049] 第八方面，提供一种通信装置，该通信装置包括处理器以及与该处理器耦合的存储器，处理器用于根据存储器中加载的程序指令执行如第二方面所述的报文传输的方法。

[0050] 第九方面，提供一种通信系统，该通信系统包括源端设备、目的端设备和至少一个路由设备，源端设备和目的端设备之间通过以太网进行远程直接内存访问RDMA，源端设备和目的端设备之间的通信路径至少包括一个路由设备相连，源端设备的网络接口卡上包括源队列对，源队列包括发送队列；目的端设备的网络接口卡上包括目的队列对，目的队列对包括接收队列；源端设备，用于从源队列对的发送队列中获取Q个数据段，分别封装该Q个数据段得到Q个报文，并分别发送该Q个报文，其中，Q个报文中的每个报文携带第一报头和第二报头，每个报文携带的第一报头用于指示本报文在目的端设备的内存中的写入地址，每个报文携带的第二报头包含源端口号信息，Q个报文中至少两个报文携带的第二报头中的源端口号信息不相同，Q为大于等于2的正整数；至少一个路由设备，用于接收源端设备发送的Q个报文，根据Q个报文中的每个报文携带的源端口号信息，分别为每个报文确定转发路径，并分别根据确定的转发路径转发每个报文；目的端设备，用于接收Q个报文，根据Q个报文各自携带的本报文在目的端设备的内存中的写入地址，分别将Q个报文从目的队列对保存到目的端设备的内存中。

[0051] 所述源端设备还用于进一步执行上述第一方面的方法，所述目的端设备还用于执行上述第二方面的方法。

[0052] 第十方面，提供一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得计算机执行如第一方面所述的报文传输的方法。

[0053] 第十一方面，提供一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得计算机执行如第二方面所述的报文传输的方法。

附图说明

[0054] 图1为本申请的实施例中数据通信系统的组成示意图。

[0055] 图2为采取RoCE协议进行传输的数据通信系统的示意图。

[0056] 图3为现有技术中两台服务器之间进行RoCE协议下的报文传输而产生负载不均衡

的示意图。

[0057] 图4为本申请的实施例中两台服务器之间进行RoCE协议下的报文传输的系统的组成示意图。

[0058] 图5为本申请的一个实施例中的源端的流程示意图。

[0059] 图6为本申请的一个实施例中的目的端的流程示意图。

[0060] 图7为现有技术中RoCEv2协议下报文的帧结构示意图。

[0061] 图8为本申请的一个实施例中封装后的报文的帧结构示意图。

[0062] 图9为本申请的一个实施例中的位图结构的示意图。

[0063] 图10为本申请的一个实施例中的位图在数据通信系统中的应用的示意图。

[0064] 图11为本申请的一个实施例中的位图在目的端接收到乱序报文时的示意图。

[0065] 图12为本申请的一个实施例中的位图在目的端接收到当前预备接收的下一个报文时的示意图。

[0066] 图13为本申请的另一个实施例中的源端的流程示意图。

[0067] 图14为本申请的另一个实施例中的目的端的流程示意图。

[0068] 图15为本申请的实施例中的源端设备的网络接口卡的功能结构的示意图。

[0069] 图16为本申请的实施例中的目的端设备的网络接口卡的功能结构的示意图。

[0070] 图17为本申请的实施例中的通信装置的结构的示意图。

[0071] 图18为本申请的实施例中的源端设备的结构示意图。

[0072] 图19为本申请的实施例中的目的端设备的结构示意图。

具体实施方式

[0073] 为了使本申请的上述目的、技术方案和优点更易于理解，下文提供了详细的描述。所述详细的描述通过使用方框图、流程图和/或示例提出了设备和/或过程的各种实施例。由于这些方框图、流程图和/或示例包含一个或多个功能和/或操作，所以本领域内人员将理解可以通过许多硬件、软件、固件或它们的任意组合单独和/或共同实施这些方框图、流程图或示例内的每个功能和/或操作。

[0074] 以下为本申请文件中相关的术语：

[0075] RDMA(Remote Direct Memory Access)技术全称远程直接内存访问，是为了解决网络传输中服务器端数据处理的延迟而产生的。RDMA通过网络把一个服务器中的数据直接传入另一个服务器的存储区，将数据从一个系统快速移动到其它系统的存储器中，而不对本系统的操作系统造成影响，这样就不需要使用到多少本系统的计算处理功能。它消除了外部存储器复制和上下文切换的开销，因而能解放内存带宽和CPU周期用于改进应用系统性能。其中，采用以太网进行RDMA被称为RoCE。

[0076] 如图1所示，在数据通信系统100，服务器可大致分为软件层和硬件层(图1中以两个服务器为例示出)，其中，软件层包括至少一个应用程序，而硬件层则主要由处理器111、内存121和网络接口卡131等组成。本实施例中，服务器101上的一个应用程序的数据需要通过RoCE协议共享到另一个服务器102上，以供另一个服务器102上的应用程序使用。

[0077] 如图2所示，数据通信系统200中包括服务器201和服务器202，其中，服务器201中包含网络接口卡241和主处理系统281，主处理系统281包括主机CPU261以及主机内存271

(其它计算机系统的常规硬件如硬盘、总线等图2未示出),主处理系统281上还运行各种软件组件,例如操作系统251,以及在操作系统251上运行的应用程序211。服务器202中包含网络接口卡242和主处理系统282,主处理系统282包括主机CPU262以及主机内存272,主处理系统282上还运行各种软件组件,例如操作系统252,以及在操作系统252上运行的应用程序212。

[0078] 网络接口卡241(也可以称为网络适配器或通信适配器)中有缓存221,缓存221中可以设置队列对(英文全称:Queue Pair,缩写:QP),图2中所示为QP231(网络接口卡中的QP根据上层应用的需求设置,可以设置多个QP,图2中以一个QP为例)。QP是网络接口卡提供给应用程序的虚拟接口,由一个发送工作队列(英文:Send Work Queue)和接收工作队列(英文:Receive Work Queue)组成,发送工作队列和接收工作队列永远是一同产生并成对出现的,它们将在其存在的时间内一直保持成对的状态。应用程序向网络接口卡发送的指令被称为工作队列元素(英文全称:Work Queue Element,WQE)。在服务器201中的应用程序211通过RDMA的方式向服务器202中的应用程序212发送数据之前,服务器201和服务器202先建立QP配对,即明确由QP231与QP232共同实现应用程序211与应用程序212之间的数据传输,并在之后发送的报文中加入相应的队列对标识(英文:QP ID)。

[0079] RDMA的工作过程通常分为三个部分。第一步,当服务器201上的应用程序211执行RDMA请求时,不执行在主处理系统的内存上的任何数据复制,RDMA请求从应用程序211的缓存被发送至网络接口卡241中的缓存221中的队列对的发送队列。第二步,网络接口卡241读取缓存221中发送队列的内容(数据),将内容通过报文的形式发送到服务器202中的QP232中,从而写入网络接口卡242的缓存222中。第三步,网络接口卡242收到数据后,直接将该数据写入主处理系统的应用程序212对应的内存中。

[0080] 在报文经过多路径的以太网从服务器201到达服务器202的过程中,以太网中的路由设备根据报文的五元组信息选择转发路径。具体来说,路由设备通过对报文的五元组信息,即报文的源端口号、目的端口号、源IP地址、目的IP地址和协议类型等五个部分进行哈希计算,根据计算出的哈希值作为报文转发路径的依据。如图3所示,该数据通信系统300中的两台服务器,服务器301和服务器302之间通过多台路由器连接在一起,并进行RoCE协议下的通信。每个服务器上有多个QP,例如,如图中所示,服务器301上有QP351和QP352,服务器302上有QP353和QP354。由于在现有技术中,同一个QP在发送数据时使用同一个源端口号,服务器301中的QP351向服务器302中的QP发送数据时,报文的五元组信息保持不变,所以作为选路依据的哈希值也相同,导致QP351发送的所有数据都会选择同样的路径,例如都选择经路由器321将数据发送至服务器302中的QP上。当QP351所发送的数据量较大时,会导致连接路由器321的网络路径的负载较大,从而使得整个报文传输系统的路径的负载不均衡。再加上RoCE网络的快启动特性,即在RoCE网络中,源服务器从启动数据发送开始,就以最大能力进行数据的发送。当网络流量达到一定值时,会导致网络产生拥塞的概率大幅增加。并且,数据通信系统300中往往不止两台服务器,可能会有更多的服务器与路由器321相连。当与路由器321连接的网络路径产生拥塞时,会影响到所有与其相连的服务器的报文传输。当网络产生拥塞之后,不仅会导致网络的时延增加,也会导致网络丢包的可能性增加。而RoCE网络对于丢包比较敏感,随着网络丢包率的增大,网络传输的有效带宽就会快速下降。

[0081] 基于希望达到RoCE协议的报文在多路径网络中的传输进一步均衡的目的，本申请提供一种粒度更细化的报文传输方法和相关的装置，在源端侧发送多个报文时，对相同QP所发出的的报文进行进一步的分组，使得属于相同QP的不同分组的报文，其报文中携带的源端口号信息不相同，从而使得由相同QP发出的报文，在经过多路径网络时，经过哈希算法获得了不同的路径，从而即使由该QP发出的报文在某个时间段内发生流量异常增大的情况，也能够避免这些流量都经过相同的路径，并避免某条路径的拥塞所引发的整个多路径网络的传输不均衡和拥塞现象。由于在源端发送报文时，本申请将相同QP发出的报文的源端口号信息进行修改，这种修改使得携带不同源端口号信息的报文在多路径网络中可能经过不同的路径到达目的端，由于各条路径的长短和效率不同，因此可能使得这些报文到达目的端的顺序与这些报文在源端发送的顺序不同，这样有可能导致目的端接收到报文后无法将这些报文保存到真正的目的地。在现有技术中，相同QP发出的报文携带相同的源端口号信息，这些报文在相同的路径进行转发，目的端的接收顺序与源端的发送顺序一致，因此RoCEv2协议规定源端相同QP发出的报文只在首包报文中携带报文中的数据在目的端的内存中的写入地址，其它非首包报文不需要携带相关的写入地址，目的端根据接收到的报文的顺序可以实现报文写入对应的内存地址。在本申请源端相同QP发出的报文采用不同的源端口号信息发出之后，为了避免乱序的报文在目的端无法写入真正的目的地的问题，本申请还对相同QP发出的报文进行扩展，使得报文的报头与现有技术存在一定区别，并由此解决对应的问题。

[0082] 图4所示的是本申请的实施例的系统结构图。如图所示，数据通信系统400包括2台服务器，服务器401和服务器402(图中所示为2台，实际数量可能是2台或更多台)。服务器401和服务器402分别与路由器411和路由器412直接相连，而路由器411和路由器412之间又通过路由器421、路由器422、路由器423、路由器424等4台路由器相连。服务器501中包括处理器431和网络接口卡441，网络接口卡441中包含若干QP，图中所示为QP451和QP452，其中每个QP对应设置有一个位图。服务器402的构成情况和服务器401相似，由处理器432和网络接口卡442组成。其中，网络接口卡441和网络接口卡442支持RoCEv2协议，服务器401和服务器402之间通过QP进行RDMA通信。在图4中的位图是本申请的其中一个实施例在目的端进行报文的接收和排序的具体实现，在其他的实施例中也可以采用其他的方法来实现。

[0083] 图5和图6所示的是本申请的一个实施例的源服务器发送报文和目的服务器接收报文的流程图。

[0084] 如图5所示，源服务器所进行的步骤如下：

[0085] S1:网络接口卡441从QP451的发送队列中获取待发送的Q个数据段。一般来说，当源服务器401中的应用程序提交一个工作请求后，该工作请求将直接被发送至网络接口卡441中相应的QP上。网络接口卡441进而可以读取该工作请求，并使QP来执行该工作请求。在本实施例中，该工作请求的内容是发送一组应用数据，该组应用数据可以包括Q个数据段，其中Q为大于等于2的正整数。

[0086] S2:确认该获取的数据段封装而成的报文将要写入目的服务器402的内存的地址。其中，该地址是根据Q个数据段的基地址和Q个数据段在该获取的数据段之前的数据段的长度计算出来的。

[0087] 在源服务器401通过RDMA的方式向目的服务器402发送数据之前，源服务器401先

和目的服务器402进行通信，目的服务器402将源服务器401将要发送的数据封装而成的报文的基地址通知源服务器401，其中，基地址指的是该组报文的第一个报文在目的服务器内存中的写入地址的首地址。

[0088] S3:封装该获取的数据段，得到封装后的报文。

[0089] 图7所示的是现有的RoCEv2的报文格式。与RoCEv1的格式相比，RoCEv2的报文格式加入了用户数据报协议(英文全称:User Datagram Protocol，缩写:UDP)报头部分，从而支持以太网IP路由功能，增强了RoCE网络的扩展性。其中，UDP的报头由源端口号、目的端口号、长度、校验和以及数据等五个部分组成。其中，在RoCEv2的报文中，UDP的目的端口号的值根据协议规定，固定为4791。由于数据通信系统中有多个服务器，每个服务器上有多个QP，每个QP的源端口号的值一般不同。

[0090] 在本申请的实施例中，对封装后的报文的扩展主要分为两部分，具体如下：

[0091] 如图8所示，一部分是给数据段增加第一报头，该第一报头中携带有用于指示本报文在目的端的内存中的写入地址的信息。具体来说，如果该数据段是该组数据中的第一个数据段，则在数据段的基本传输报头(英文全称:Base Transport Header，缩写:BTH)部分后加入RDMA扩展传输报头(英文全称:RDMA Extended Transport Header，缩写RETH)部分；如果不是第一个数据段，则在数据段的BTH部分后加入扩展报头(英文全称:Extended Header，缩写:EXH)部分。其中，每个WQE的第一个和最后一个数据段的BTH部分分别包含相应的信息，用以指示该报文是WQE中的第一个数据段或者是最后一个数据段。

[0092] 其中，RETH部分包括虚拟地址(英文:Virtual Address)、远程秘钥(英文:Remote Key)、DMA长度(英文:DMA Length)等三个部分。其中Virtual Address部分的长度是64比特，记载的是进行RDMA操作后对应的目的端的虚拟地址;Remote Key部分的长度是32比特，记载的是允许进行RDMA操作的授权信息;DMA Length部分的长度是32比特，记载的是进行DMA操作的报文的字节数。EXH头部包括Virtual Address、立即数(英文:Immediate)、WQE序号(英文:WQE Number)、保留字段(英文:Reserved)等四个部分。其中，Virtual Address部分和RETH头部中的Virtual Address部分一样，长度是64比特，记载的是当前报文需要写入的目的端的内存地址;Immediate部分的长度是1比特，记载的是当前报文是否携带立即数;WQE Number部分的长度是31比特，记载的是QP发送的WQE序号;Reserved部分的长度是32比特，为保留字段。其中，EXH头部除了必须要有Virtual Address部分外，剩下的三个部分可以根据实际需要进行调整。

[0093] 采用将包含虚拟地址的报头封装在报文上的做法，可以使得报文到达目的端时可以快速地写入内存之中。同时，由于报文具有虚拟地址部分，即使报文在网络传输途中发生了乱序，也可以通过根据虚拟地址写入目的端的内存中的相应位置。

[0094] 另一部分是给数据段增加第二报头。其中，第二报头中携带源队列对的源端口号信息。和现有技术相比，本申请的实施例中的Q个数据段封装而成的Q个报文中至少有两个报文具有不同的源端口号信息。路由器在根据五元组信息选择转发的路径时，由于其中的源端口信息不同，具有不同源端口号信息的报文很大可能会选择不同的转发路径。因为这种对同一QP发送的报文的不同的源端口号信息的设置方式，能够将同一QP发出的报文流量分担到不同的转发路径，即使该QP发出的报文流量出现较大的情况，也不会引起整个多路径网络的某条路径的拥塞。

[0095] 可选的，还可以在数据段中的BTH部分添加报文序号(英文全称:Packet Sequence Number，缩写:PSN)，该报文序号用来表示该数据段在Q个数据段中的顺序。

[0096] S4:每当一个数据段封装成报文后，发送该报文。

[0097] S5:判断是否已经发送预设数量的报文。当已经发送预设数量的报文后，进行S6；当还没有发送预设数量的报文后，跳转至S1。

[0098] 可选的，判断是否已经发送预设数量的报文，并对源队列对的端口信息进行更新时，该预设数量可以是变化的。例如，可以先发送3个报文后，对源队列对的端口信息进行更新，再发送4个报文后，对源队列对的端口信息进行更新。该预设数量也可以是固定的，例如每次都在发送3个报文后，对源队列对的端口信息进行更新。

[0099] S6:当已经发送预设数量的报文后，对源队列对的端口信息进行更新。通过这种方法，使得该组数据封装而成的报文的第二报头，具有不同的源端口号。从而当报文在网络中进行传输时，路由器根据报文的五元组信息的哈希值进行选路。由于报文拥有不同的源端口号，所以得出的哈希值很可能会不相同，从而选择不同的路径进行传输，使得网络中各个路径的流量更加均衡。

[0100] 对于当前RoCE协议的规定而言，每个QP只会采用固定的源端口号，其在网络中的转发路径是固定的，只要不发生丢包的情况，报文就不会出现乱序的问题。而在上述实施例中，为了达到流量均衡的目的，QP对应的源端口号信息是在发生变化的，因此在网络上报文转发的路径也在发生着变化。由于不同网络路径对报文的处理时间可能会不同，所以报文在目的端可能会出现乱序。而源端通过给报文封装RETH或EXH扩展报头，将欲写入的目的服务器的内存的虚拟地址放入报文中。当报文到达目的端时，可以直接根据RETH或EXH扩展报头里的虚拟地址写入目的服务器相应的内存位置，相当于已经恢复了其在源端发送时的顺序。

[0101] S7:判断该Q个数据段是否已经发送完。如果还有数据段没有发送，则跳转至S1。

[0102] 需要注意的是，上述S1-S7的编号仅用来指代，并不意味着在本申请的实施例中，上述步骤需要按照特定顺序来执行。例如，S2确认写入地址的步骤也可以S1之前。

[0103] 在本申请的另一个实施例中，源端还可以将待发送的Q个数据段分成至少两个分组，每个分组包含至少一个数据段。封装每个分组中的数据段得到每个分组中的报文，其中，每个分组中的报文携带的源端口号信息相同，至少两个分组中的报文携带的源端口号信息不同。在源端发出Q个报文之后，该Q个报文经过路由器进行转发。路由器根据该Q个报文的五元组信息选择转发路径。当Q个报文的源端口号信息不同时，可能会被选择以不同的路径进行转发，因此该Q个报文到达目的端的顺序可能与该Q个报文在源端发出的顺序不同。目的端在接收到该Q个报文后，需要将该Q个报文中的数据段保存到对应的地址。另外，现有技术中，RoCE协议规定目的端按报文的发送顺序接收报文，如果接收到的报文发生乱序，目的端要马上发送重传指示，使得源端重新发送在传输路径中可能丢失的报文。然而，本申请的上述实施例在发送端改变了相同QP发出的报文的源端口号信息，使得上述Q个报文到达目的端的顺序与该Q个报文的发送顺序大概率不同，在这种情况下，如果目的端一旦确定接收的报文发生乱序，就马上发送重传指示，报文重传的代价会较大。本申请在目的端还对接收到的报文进行乱序检测，在检测到发生乱序情况时并不立刻向源端发送报文重传指示，而是启动报文丢包检测流程，在根据该丢包检测流程确定发生丢包时，才向源端发送

报文重传指示，提高系统的传输效率。目的端的具体流程实施例参见图6。

[0104] 图6中，在本申请的实施例中，通过接收到的报文携带的报文序号，用来检验源服务器发送的报文是否出现乱序、丢包以及判断是否收齐。其中，该方法可以通过位图、数组和链表等方式进行实现。本申请的实施例将以位图为例进行说明。

[0105] 图9-12所示的是位图算法在本申请的实施例中的原理，其中：

[0106] 图9所示的是实现位图算法的位图示意图。如图9所示，在本申请的实施例中，每个QP都对应一个位图，用来记录报文的接收情况。每个位图包括多个位图位，每个位图位代表一个报文，将位图的位图位从前向后进行标号，并与报文的报文序号的值建立起对应关系，每个位图位按照报文的发送顺序从前往后与报文进行对应。每个位图还具有尾指针和头指针，尾指针指向该位图对应的队列对的接收队列预备接收的下一个报文所对应的位图位；头指针指向当前最新接收到的报文对应的位图位。当位图中的位图位的值为有效时，代表着该位图位对应的报文已经收到；当位图中的位图位的值为无效时，代表着该位图位对应的报文还没有收到，其中，有效既可以用值为1来表示，也可以用值为0来表示，在本申请的实施例中，有效用值为1来表示。同时，根据所需要进行排序的报文的报文序号的值的范围，设定所使用的位图的范围，若源端发送Q个报文，则目的端对应的位图至少包括Q个位图位。在该位图的范围中，最前端对应具有最小的报文序号的值的报文。

[0107] 其中，上述的尾指针的指向是预备接收的下一个报文，一般指该位图对应的队列对的接收队列当前未接收到，并在下一个即将接收到的报文中预备接收到的报文，并且，该预备接收到的下一个报文是当前未接收到的报文中最新发送的报文，也可以说，预备接收的下一个报文一般指目的端当前没有接收到的报文中具有最小报文序号的报文，例如，源端发送Q个报文，该Q个报文的发送顺序为1到Q，报文序号为1表示最先发送的报文，报文序号为Q表示最后发送的报文，如果目的端接收到了报文序号分别为1、2、5的报文，则预备接收的下一个报文为报文序号为3的报文，该尾指针也指向报文序号为3的报文对应的位图位。

[0108] 图10-12所示的是位图中位图位的值以及头指针和尾指针的位置是如何根据接收到的报文而发生变化的。例如，如图10所示，服务器401中的QP451向服务器402中的QP452发送10个报文，报文的报文序号分别为1-10，对应的位图也有10个位图位，从前端到后端(图中所示为从右向左)分别编号为1-10，从而和报文一一对应。报文在传输过程中顺序发生了改变，到达目的端QP452的顺序为3、1、2、4、5、6、7、8、9、10。

[0109] 如图11所示，当目的端QP453接收到报文序号的值为3的报文时，头指针移动到对应的3号位图位上，并将该位图位的值置为有效。由于尾指针所指向的位置是预备接收的下一个报文所对应的位图位，也就是报文序号的值为1的报文，因此尾指针保持不动。

[0110] 如图12所示，当目的端QP453接收到报文序号的值为1的报文时，头指针移动到对应的1号位图位上，并将该位图位的值置为有效。而尾指针收到了当前预备接收的下一个报文，因此发生移动，其新的指向为当前接收到的报文所对应的位图位之后的无效的位图位中的第一个位图位，即2号位图位。

[0111] 如图6所示，在本申请的一个实施例中，目的端所进行的步骤如下：

[0112] S1：目的端依次接收源端发送的报文，将该报文缓存到相应的目标队列对中。

[0113] S2：由于接收的报文在源端时，报头被加入了RETH部分或者EXH部分，而RETH部分

和EXH部分中都包含有该报文所应该写入的目的端的内存的地址。目的端根据报文中包含的虚拟地址，将该接收到的报文写入相应的内存的地址中。

[0114] S3：每接收到所述Q个报文中的一个报文，记录当前接收到的报文携带的报文序号，并根据当前接收到的报文的报文序号，确定预备接收的下一个报文的报文序号；在接收到所述Q个报文中的下一个报文后，确定接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致。

[0115] 以采取位图进行检验的方式为例，当接收到源端发送的报文后，根据该报文的报文序号，将位图中代表该报文所对应的位图位的值设置为有效，即为1，并将该位图的头指针指向该报文对应的位图位。而位图的尾指针指向的是当前预备接收的下一个报文所对应的位图位。因此，当位图中的头指针和尾指针指向不同的位图位时，可以判断接收到的报文不是当前预备接收的下一个报文，即接收到的报文发生乱序。当接收到的报文是当前预备接收的下一个报文时，直接进行S5，判断报文是否已经收齐；如果接收到的报文不是当前预备接收的下一个报文时，进行S4。

[0116] S4：当接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号不一致时，启动丢包检测流程，判断报文在传输过程中是否发生丢包。仍然以采取位图进行检验的方式为例，当判断接收到的报文不是当前预备接收的下一个报文的情况产生时，启动定时器。如果定时器超时后，尾指针的指向没有发生改变，说明在预设的时间内，目的端没有接收到尾指针所指向的位图位所对应的报文，从而说明尾指针当前指向的位图位所对应的报文发生了丢包。而如果尾指针当前指向的位图位所对应的报文接收到了，尾指针发生移动，定时器也将进行重置。

[0117] 判断报文在传输过程中是否发生丢包还有另一种方法。当判断有接收到的报文不是当前预备接收的下一个报文的情况产生时，确定头指针当前指向的位图位是否超过了预定值T。如果头指针当前指向的位图位超过了预定值T，说明当前头指针指向的位图位和尾指针指向的位图位之间的某个位图位所对应的报文发生了丢包。该预定值T可以根据实际需要进行设定，例如，该预定值T可以设定为Q，也就是该组报文的数目，在这种情况下，当头指针指向的位图位超过了预定值T时，说明目的端在该组报文还没有收齐的情况下，已经接收到了下一组报文，可以判断发生了丢包。

[0118] 如果通过所述丢包检测流程确定在报文传输过程中发生丢包，目的端向源端发送否定应答报文，通知源端，报文的传输过程有错误。同时，向源端发送报文重传指示，该重传指示中携带有尾指针当前指向的位图位所对应的报文的报文序号，以请求源端重传该报文序号所对应的报文之后的所有报文。通过这种做法，在目的端接收到乱序的报文时，可以较为准确的确定哪些情况可能发生丢包，并在确定发生丢包时，才指示源端进行报文重传，从而提高了系统的效率。

[0119] 当判断没有出现丢包时，进行S5。

[0120] S5：判断报文是否已经收齐。当该组报文所对应的位图位的值都被置为有效时，说明该组报文已经被收齐，进行S6。如果报文还没有被收齐，则重新回到S1。

[0121] S6：当报文已经被收齐时，目的端向源端发送确认应答报文。

[0122] 图13和图14所示的是本申请的另一个实施例的源端和目的端的流程图。

[0123] 如图13所示，源端所进行的步骤如下：

[0124] S1:网络接口卡441从QP451的发送队列中获取待发送的Q个数据段。

[0125] S2:封装该获取的数据段,得到封装后的报文。和本申请的前述的实施例不同,在这里只向数据段添加携带该源队列对的端口信息的第二报头,以及给每组数据的第一个数据段添加携带写入目的端内存的RETH报头,而不向剩下的数据段添加携带写入目的端内存的EXH报头。

[0126] S3:每当一个数据段封装成报文后,发送该报文。

[0127] S4:判断是否已经发送预设数量的报文。当已经发送预设数量的报文后,进行S5;当还没有发送预设数量的报文后,跳转至S1。

[0128] S5:当已经发送预设数量的报文后,对源队列对的端口信息进行更新。

[0129] S6:判断该组数据是否已经发送完。如果还有数据没有封装并发送,则跳转至S1。

[0130] 如图14所示,在本申请的第二个实施例中,目的端所进行的步骤如下:

[0131] S1:目的端依次接收源端发送的报文,并将这些报文缓存至相应的队列对中。

[0132] S2:判断接收到的报文是否是当前预备接收的下一个报文。如果不是,进行S3;如果报文是,则进行S4。

[0133] S3:判断报文是否发生丢包。如果报文发生丢包,目的端向源端发送否定应答报文,通知源端,报文的传输过程有错误,并向源端发送报文重传指示。如果报文没有发生丢包,则进行S4。

[0134] S4:判断报文是否已经收齐。如果报文已经收齐,进行S5;如果报文还没有收齐,则重新回到S1。

[0135] S5:当报文已经收齐后,根据报文所携带的报文序号进行乱序重排,使得缓存中的报文恢复顺序。

[0136] S6:当缓存中的报文被排好序后,将其写入内存之中。

[0137] S7:目的端向源端发送确认应答报文。

[0138] 基于上述技术方案,参阅图15所示,本申请的实施例提供一种网络接口卡1500,该网络接口卡1500位于位于远程直接内存访问RDMA的源端设备,该网络接口卡1500上设置有源队列对,源队列对包括发送队列;该网络接口卡1500包括:

[0139] 获取模块1510,用于从至少两个源队列对中的第一源队列对的发送队列中获取Q个数据段;

[0140] 发送模块1520,用于封装Q个数据段得到Q个报文,并发送Q个报文,其中,Q个报文中的每个报文携带第一报头和第二报头,每个报文携带的第一报头用于指示本报文在目的端设备的内存中的写入地址,每个报文携带的第二报头包含源端口号信息,Q个报文中至少两个报文携带的第二报头中的源端口号信息不相同,Q为大于等于2的正整数,目的端设备为RDMA的目的端设备;

[0141] 确定模块1530,用于根据Q个数据段的第一个数据段的基地址和每个数据段的长度,确定Q个报文中的每个报文在目的端设备的内存中的写入地址。

[0142] 本申请的实施例所提供的网络接口卡1500,其功能的实现可以参考如图5所示的报文传输的方法。

[0143] 基于上述技术方案,参阅图16所示,本申请的实施例提供另一种网络接口卡1600,该网络接口卡1600位于远程直接内存访问RDMA的目的端设备,该网络接口卡1600上设置有

目的队列对,目的队列对包括接收队列;该网络接口卡1600包括:

[0144] 接收模块1610,用于接收Q个报文,其中,每个报文携带第一报头和第二报头,每个报文携带的第一报头用于指示本报文在目的端设备的内存中的写入地址,每个报文携带的第二报头包含源端口号信息,Q个报文中至少两个报文携带的第二报头中的源端口号信息不相同,Q为大于等于2的正整数;

[0145] 执行模块1620,用于根据Q个报文各自携带的本报文在目的端设备的内存中的写入地址,分别将Q个报文从目的队列对保存到目的端设备的内存中。

[0146] 检测模块1630,在接收模块每接收到一个报文,检测模块用于记录当前接收到的报文携带的报文序号,并根据当前接收到的报文的报文序号,确定预备接收的下一个报文的报文序号;在接收到下一个报文后,确定接收到的下一个报文的报文序号是否与预备接收的下一个报文的报文序号是否一致,如果否,启动丢包检测流程;如果通过丢包检测流程确定在报文传输过程中发生丢包,则向源端设备发送报文重传指示。

[0147] 本申请的实施例所提供的网络接口卡1600,其功能的实现可以参考如图6所示的报文传输的方法。

[0148] 图17为依据本申请的实施例的通信装置1700的结构示意图。本实施例中的通信装置可以是上述各实施例中的网络接口卡的其中一种具体实现方式。

[0149] 如图17所示,通信装置包括处理器1701,处理器1701与存储器1705连接。处理器1701可以为中央处理单元CPU,或现场可编程门阵列(英文全称:Field Programmable Gate Array,缩写:FPGA),或数字信号处理器(英文全称:Digital Signal Processor,缩写:DSP)等计算逻辑或以上任意计算逻辑的组合。处理器1701也可以为单核处理器或多核处理器。

[0150] 存储器1705可以是RAM存储器、闪存、ROM存储器、EPROM存储器、EEPROM存储器、寄存器、硬盘、移动硬盘、CD-ROM或者本领域熟知的任何其它形式的存储介质,存储器可以用于存储程序指令,该程序指令被处理器1701执行时,处理器执行上述实施例中的源端或目的端的方法。

[0151] 连接线1709用于在通信装置的各部件之间传递信息,连接线1709可以使用有线的连接方式或采用无线的连接方式,本申请并不对此进行限定。连接1709还连接有网络接口1704。

[0152] 网络接口1704使用例如但不限于电缆或电绞线一类的连接装置,来实现与其他设备或网络1711之间的通信,网络接口1704还可以通过无线的形式与网络1711互连。

[0153] 本申请实施例的一些特征可以由处理器1701执行存储器1705中的程序指令或者软件代码来完成/支持。存储器1705上在加载的软件组件可以从功能或者逻辑上进行概括,例如,图15所示的获取模块、发送模块等功能/逻辑模块,或者图16所示的接收模块和执行模块等功能/逻辑模块等。

[0154] 在本申请的一个实施例中,当存储器1705加载进程序指令后,处理器1701执行存储器中的上述功能/逻辑模块相关的事务。

[0155] 此外,图17仅仅是一个通信装置的例子,通信装置可能包含相比于图17展示的更多或者更少的组件,或者有不同的组件配置方式。同时,图17中展示的各种组件可以用硬件、软件或者硬件与软件的结合方式实施,例如,该通信装置可以以一个芯片的形式来实现。在这种情况下,存储器和处理器可以在一个模块中实现,存储器中的指令可以是预先写

入所述存储器的，也可以是后续处理器在执行的过程中加载的。

[0156] 本申请的实施例提供一种设备，如图18所示，该设备1800包括主处理系统1810和网络接口卡1830。其中，主处理系统1810用于处理业务，在需要将业务数据发送到目的端设备时，将业务数据发送到网络接口卡1830中的所述业务数据对应的源队列对的发送队列；网络接口卡1830，用于从业务数据对应的源队列对的发送队列中获取Q个数据段，该Q个数据段属于业务数据，封装Q个数据段得到Q个报文，并发送该Q个报文，其中，Q个报文中的每个报文携带第一报头和第二报头，每个报文携带的第一报头用于指示本报文在目的端的内存中的写入地址，每个报文携带的第二报头包含源端口号信息，Q个报文中至少两个报文携带的第二报头中的源端口号信息不相同，Q为大于等于2的正整数。网络接口卡1830还用于根据Q个数据段的第一个数据段的基地址和每个数据段的长度，确定Q个报文中的每个报文在目的端设备的内存中的写入地址。

[0157] 本申请的实施例还提供另一种设备，如图19所示，该设备1900包括主处理系统1910和网络接口卡1930。其中，主处理系统1910用于从设备1900的内存1920中获取应用数据，以及根据该应用数据处理业务；网络接口卡1930用于接收通过Q个报文实现的应用数据以及将接收到的Q个报文写入内存1920。其中，网络接口卡1930接收Q个报文的方法可参考如图6所示的报文传输的方法。

[0158] 本申请的实施例还提供一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得计算机执行如图5所示的报文传输的方法。

[0159] 本申请的实施例还提供另一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得计算机执行如图6所示的报文传输的方法。

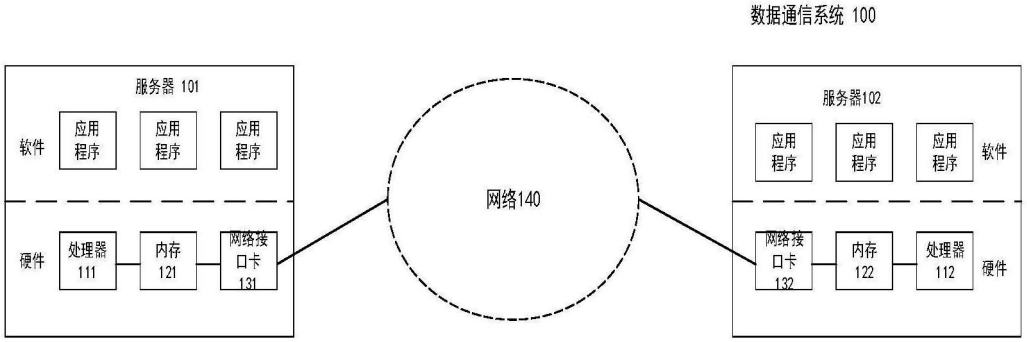

数据通信系统 100
服务器 101
软件
应用程序
应用程序
应用程序
硬件
处理器 111
内存 121
网络接口卡 131
网络140
服务器102
应用程序
应用程序
应用程序
软件
网络接口卡 132
内存 122
处理器 112
硬件

图1

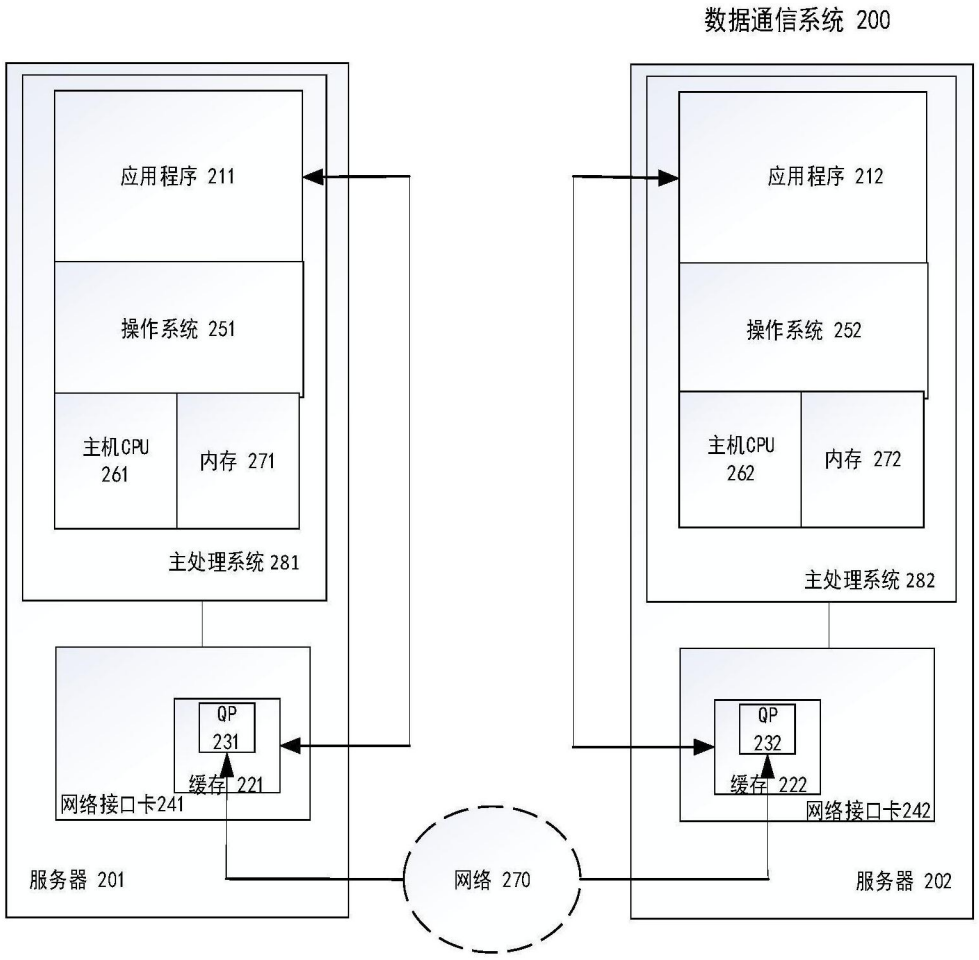

数据通信系统 200
应用程序 211
操作系统 251
主机CPU 261
内存 271
主处理系统 281
QP 231
缓存 221
网络接口卡241
服务器 201
网络 270
应用程序 212
操作系统 252
主机CPU 262
内存 272
主处理系统 282
QP 232
缓存 222
网络接口卡242
服务器 202

图2

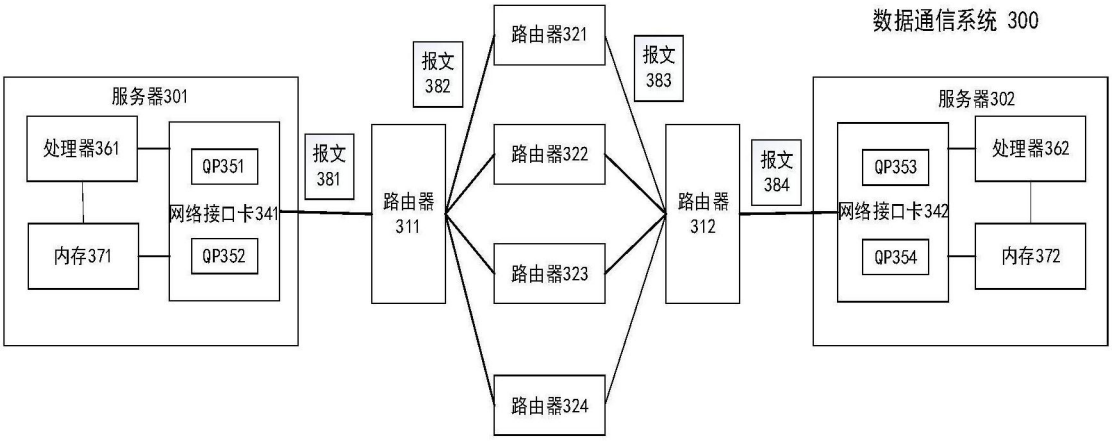

数据通信系统 300
服务器301
处理器361
内存371
QP351
QP352
网络接口卡341
报文 381
报文 382
路由器 311
路由器321
路由器322
路由器323
路由器324
报文 383
路由器 312
报文 384
服务器302
QP353
QP354
网络接口卡342
处理器362
内存372

图3

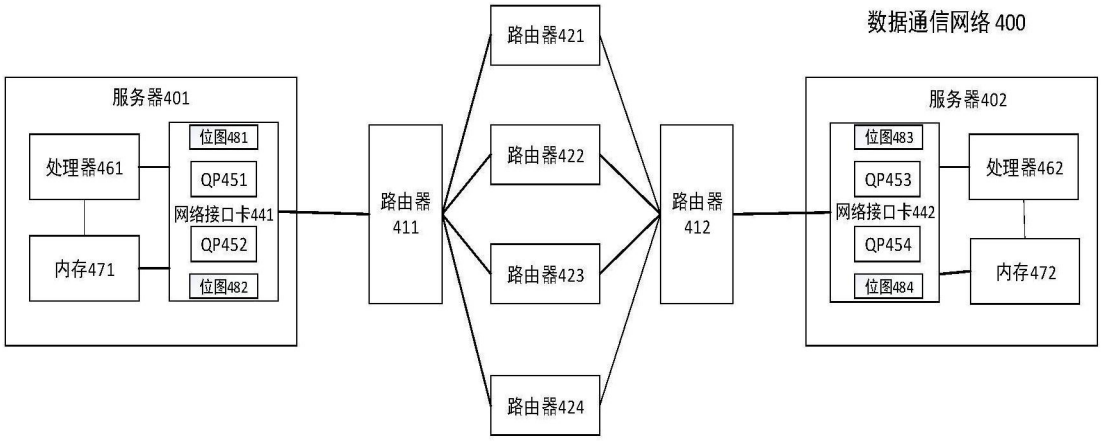

数据通信网络 400
服务器401
处理器461
内存471
位图481
QP451
QP452
位图482
网络接口卡441
路由器 411
路由器421
路由器422
路由器423
路由器424
路由器 412
服务器402
位图483
QP453
QP454
位图484
网络接口卡442
处理器462
内存472

图4

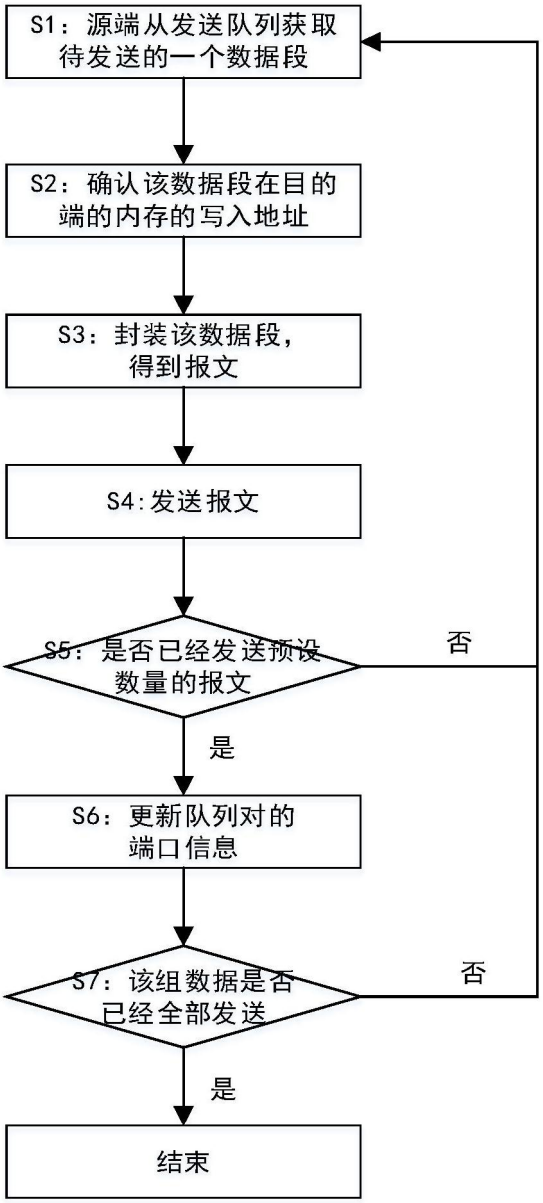

S1：源端从发送队列获取待发送的一个数据段
S2：确认该数据段在目的端的内存的写入地址
S3：封装该数据段，得到报文
S4:发送报文
S5：是否已经发送预设数量的报文
否
是
S6：更新队列对的端口信息
S7：该组数据是否已经全部发送
否
是
结束

图5

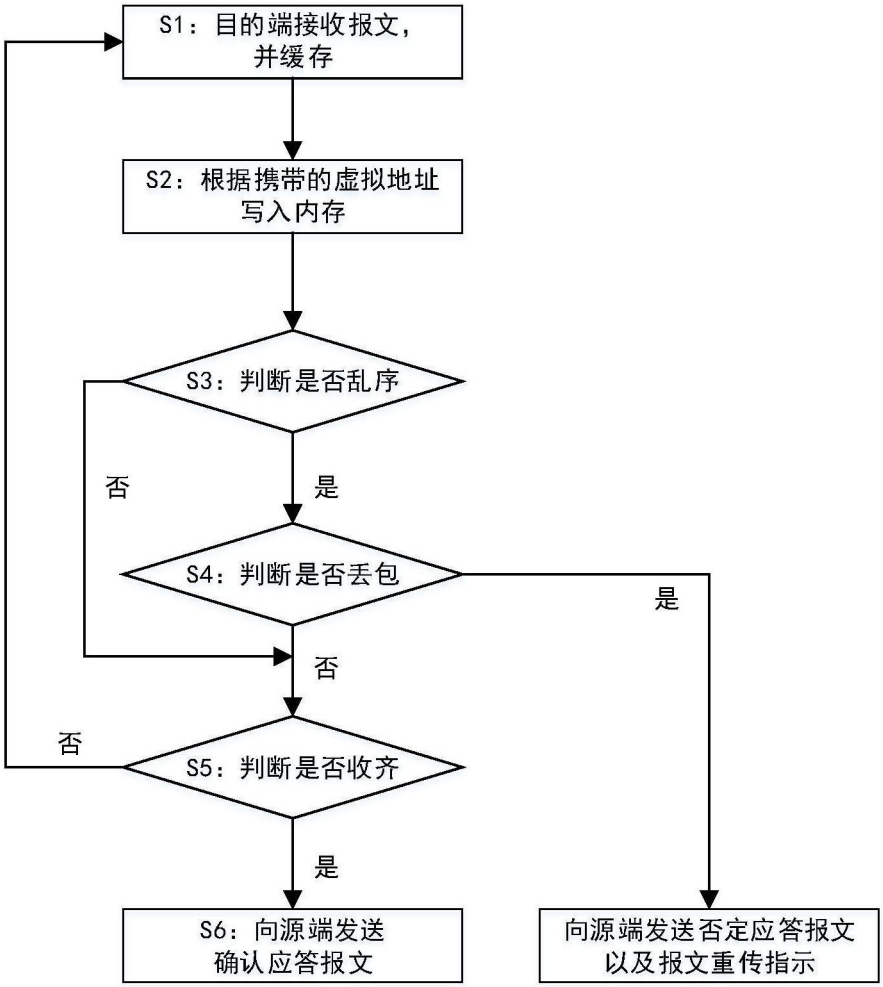

S1：目的端接收报文，
并缓存
S2：根据携带的虚拟地址
写入内存
S3：判断是否乱序
否
是
S4：判断是否丢包
是
否
否
S5：判断是否收齐
是
S6：向源端发送
确认应答报文
向源端发送否定应答报文
以及报文重传指示

图6

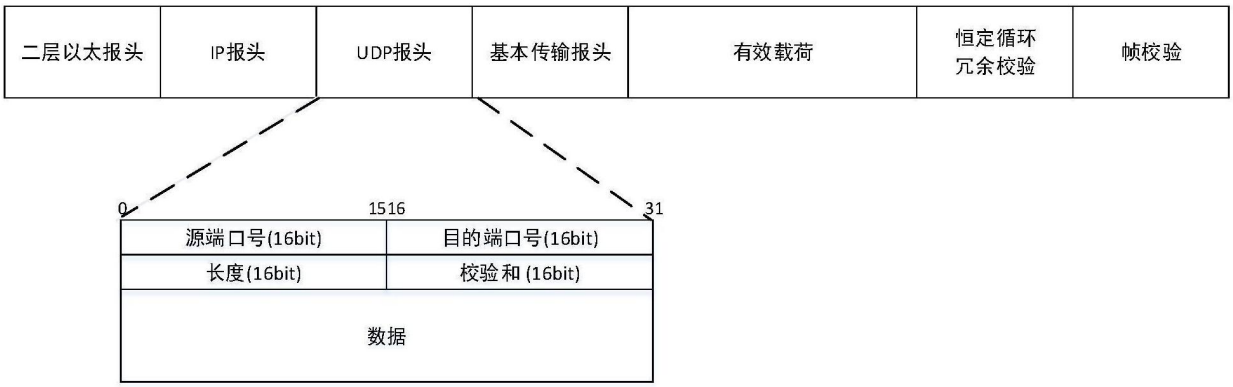

二层以太报头
IP报头
UDP报头
基本传输报头
有效载荷
恒定循环
冗余校验
帧校验
0
15
16
31
源端口号(16bit)
目的端口号(16bit)
长度(16bit)
校验和 (16bit)
数据

图7

二层以太报头	IP报头	UDP报头	基本传输报头	RDMA扩展传输报头/扩展报头	有效载荷	恒定循环冗余校验	帧校验

图8

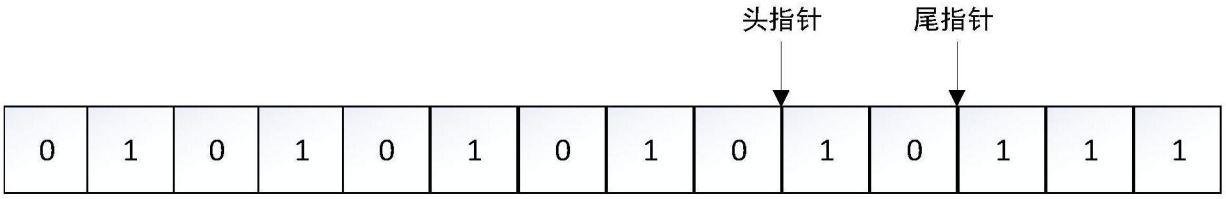


图9

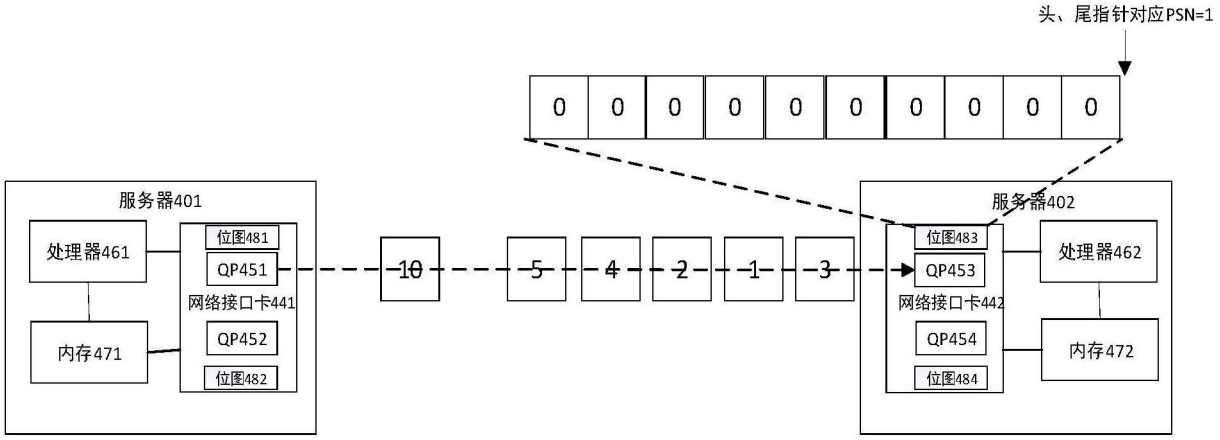


图10

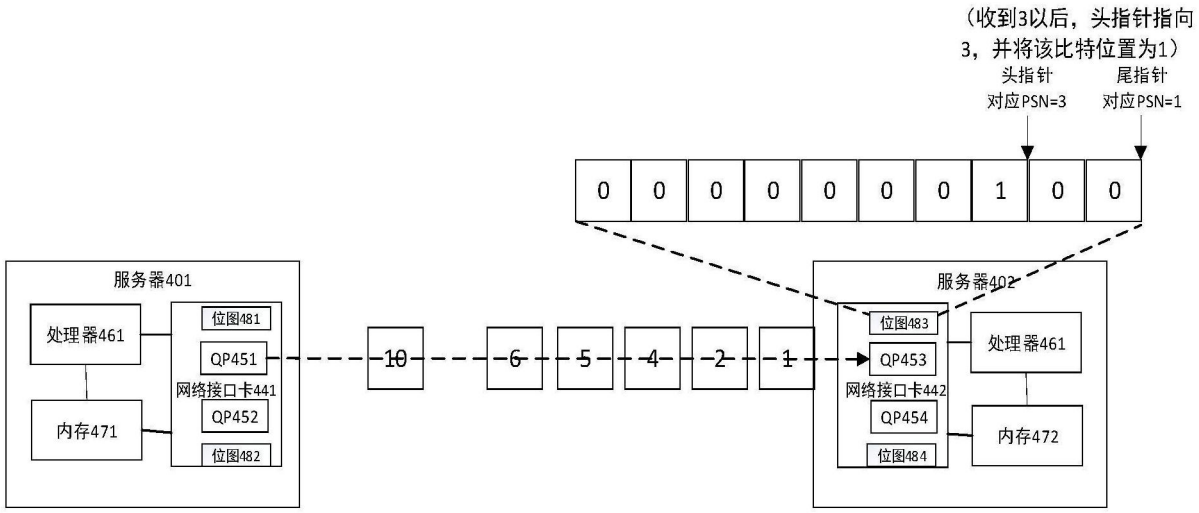

（收到3以后，头指针指向3，并将该比特位置为1）
头指针
对应PSN=3
尾指针
对应PSN=1
0 0 0 0 0 0 0 1 0 0
服务器401
处理器461
内存471
位图481
QP451
网络接口卡441
QP452
位图482
10 6 5 4 2 1
服务器402
位图483
QP453
网络接口卡442
QP454
位图484
处理器461
内存472

图11

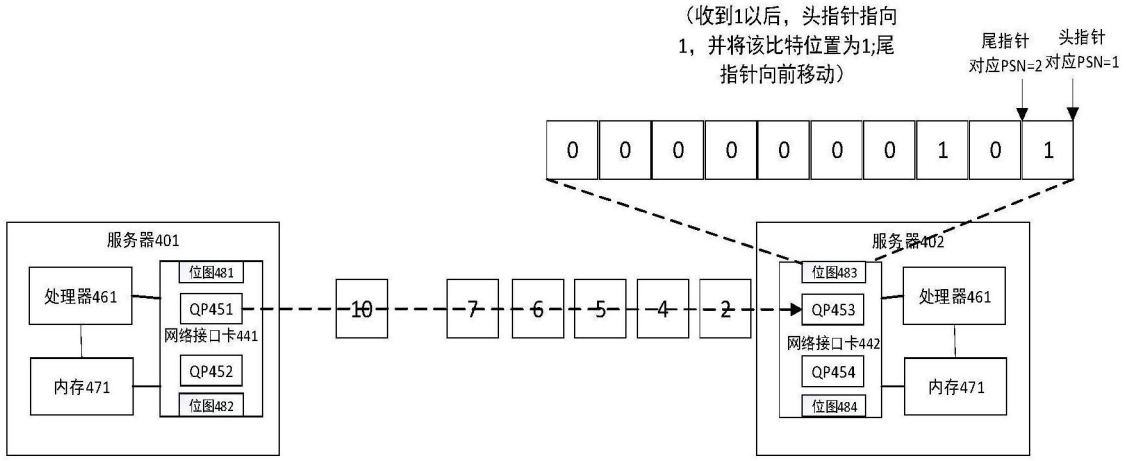

（收到1以后，头指针指向1，并将该比特位置为1;尾指针向前移动）
尾指针
对应PSN=2
头指针
对应PSN=1
0 0 0 0 0 0 0 1 0 1
服务器401
处理器461
内存471
位图481
QP451
网络接口卡441
QP452
位图482
10 7 6 5 4 2
服务器402
位图483
QP453
网络接口卡442
QP454
位图484
处理器461
内存471

图12

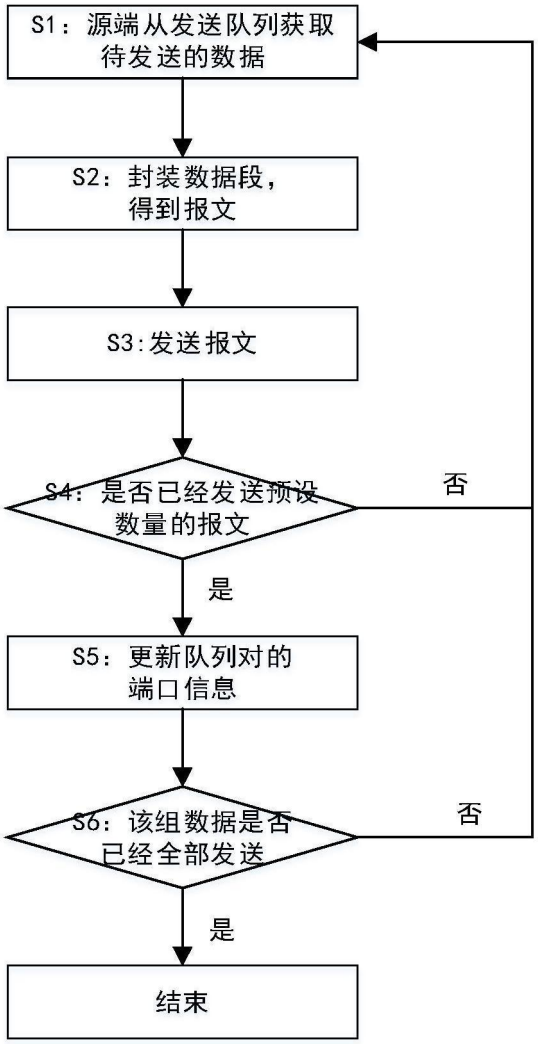

S1：源端从发送队列获取待发送的数据
S2：封装数据段，得到报文
S3:发送报文
S4：是否已经发送预设数量的报文
否
是
S5：更新队列对的端口信息
S6：该组数据是否已经全部发送
否
是
结束

图13

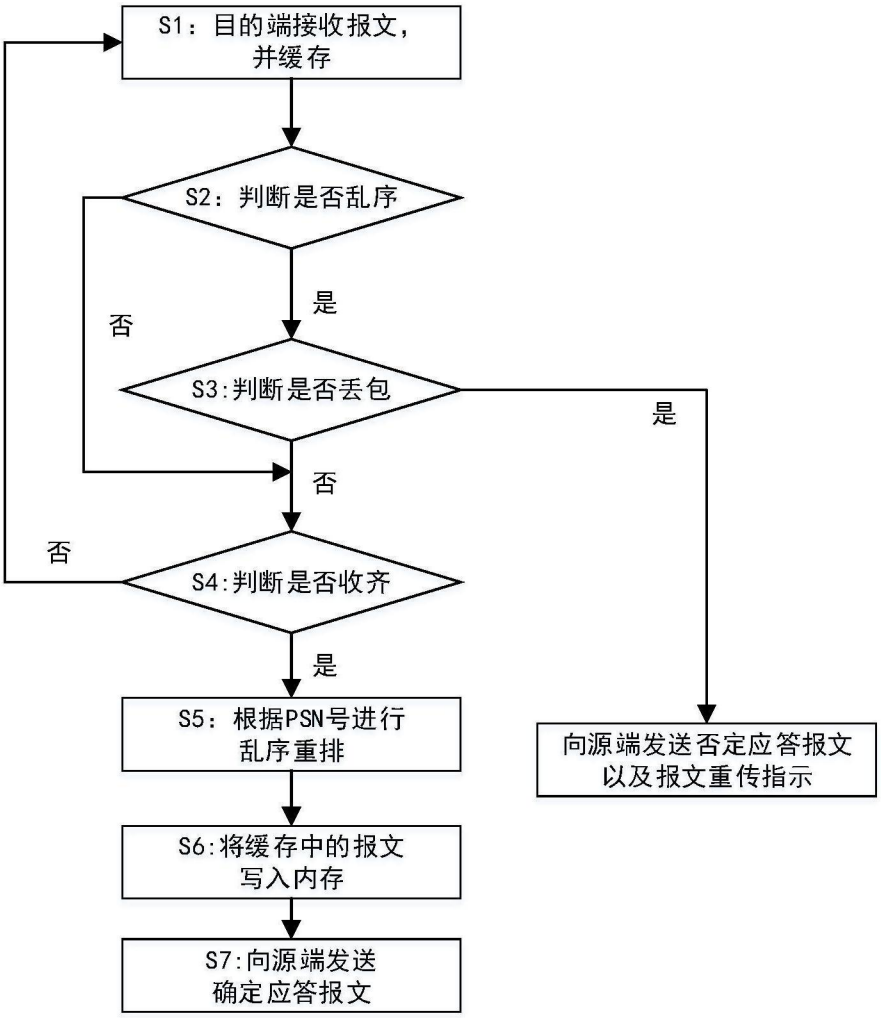

S1：目的端接收报文，
并缓存
S2：判断是否乱序
是
否
S3:判断是否丢包
是
否
否
S4:判断是否收齐
是
S5：根据PSN号进行
乱序重排
向源端发送否定应答报文
以及报文重传指示
S6:将缓存中的报文
写入内存
S7:向源端发送
确定应答报文

图14

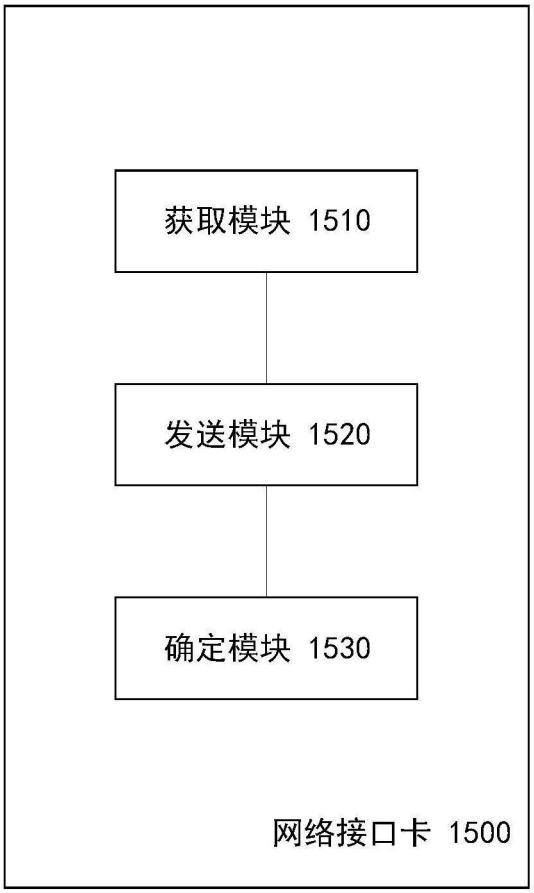


图15

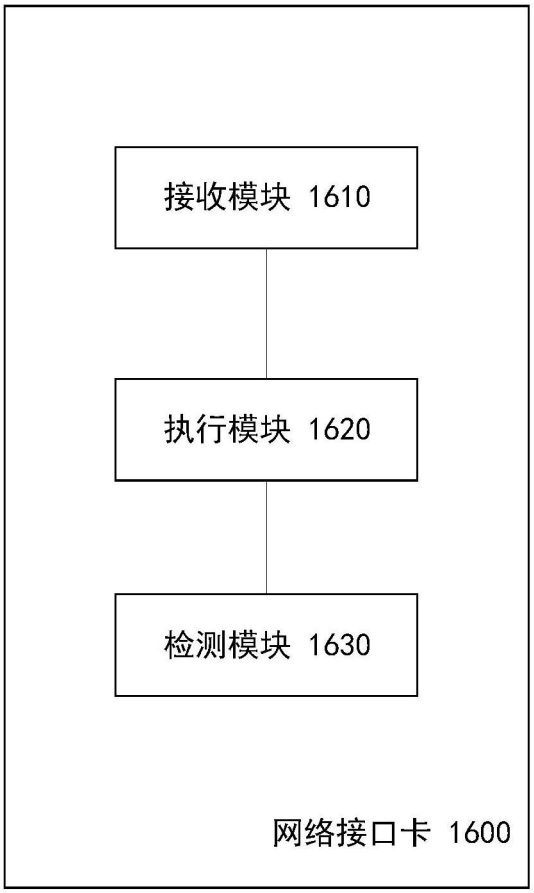

接收模块 1610
执行模块 1620
检测模块 1630
网络接口卡 1600

图16

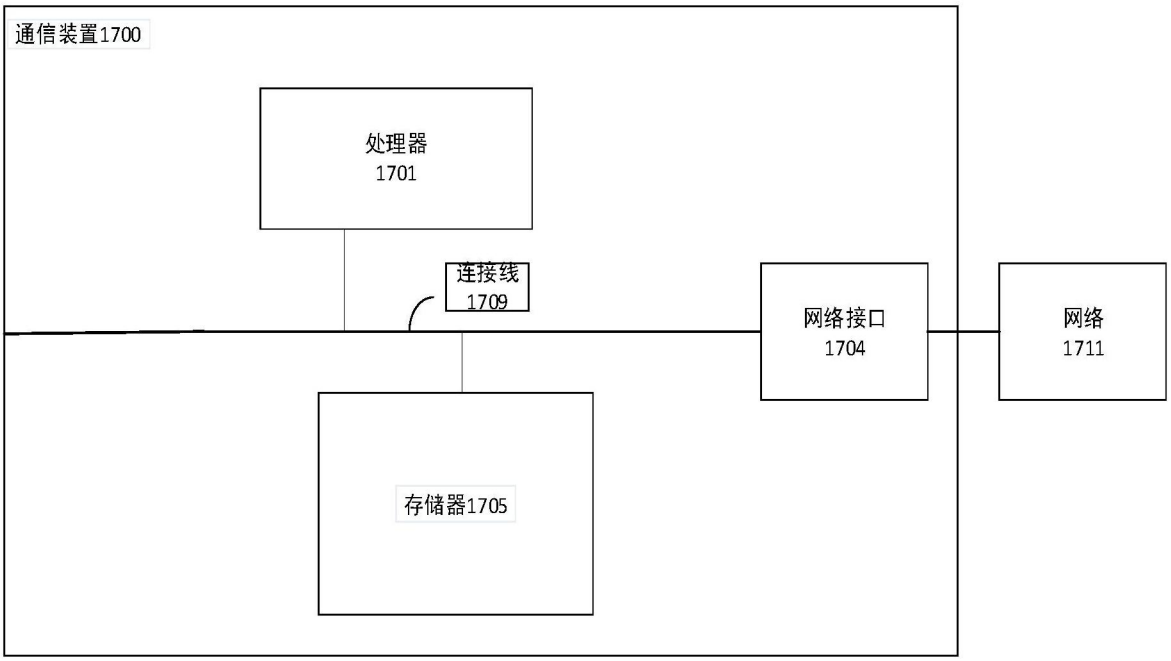

通信装置1700
处理器
1701
连接线
1709
网络接口
1704
网络
1711
存储器1705

图17

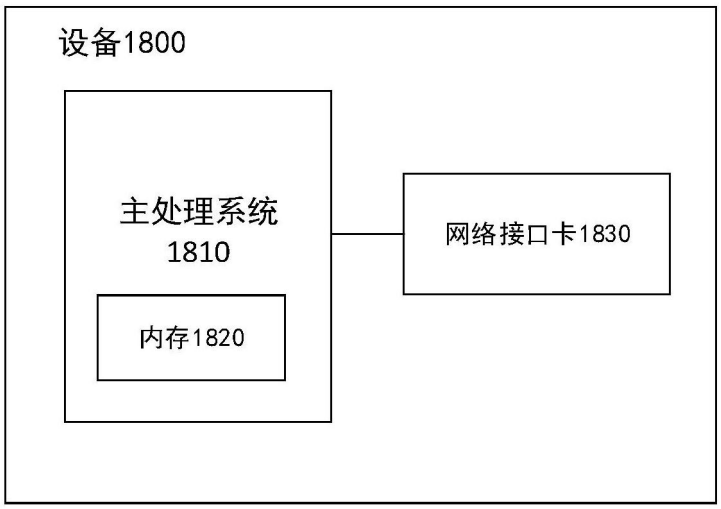

设备1800
主处理系统
1810
内存1820
网络接口卡1830

图18

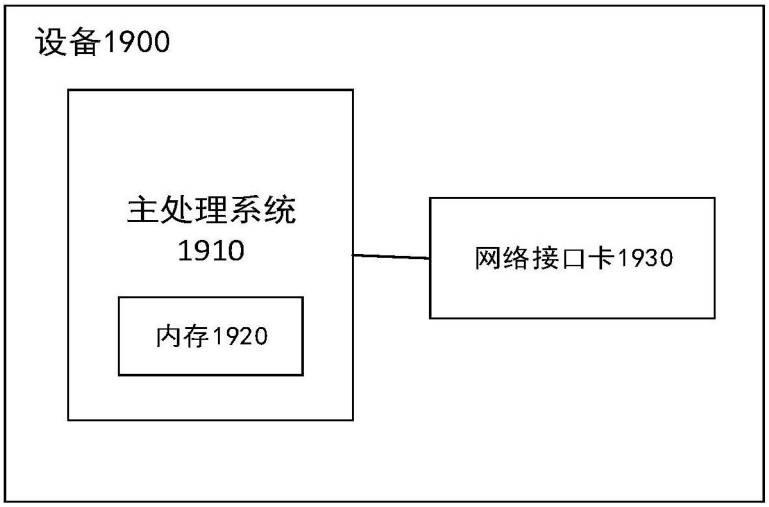

设备1900
主处理系统
1910
内存1920
网络接口卡1930

图19