

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号

WO 2019/057005 A1

(43) 国际公布日

2019 年 3 月 28 日 (28.03.2019)

(51) 国际专利分类号:

G06F 13/14 (2006.01)

(21) 国际申请号:

PCT/CN2018/105957

(22) 国际申请日:

2018 年 9 月 17 日 (17.09.2018)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201710867695.7 2017年9月22日 (22.09.2017) CN

(71) 申请人: 华为技术有限公司 (HUAWEI

TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东

省深圳市龙岗区坂田华为总部办公楼,

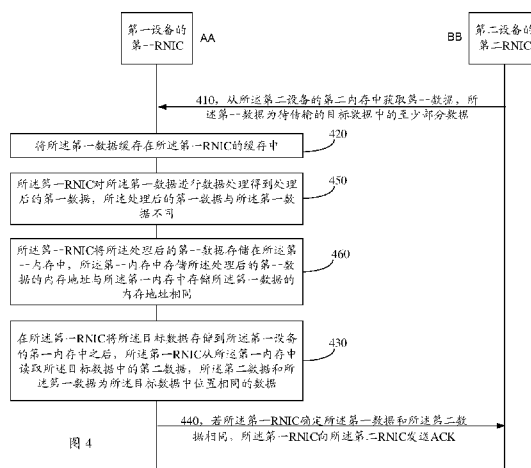
Guangdong 518129 (CN)。

(72) 发明人: 曲会春 (QU, Huichun); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 吴沛 (WU, Pei); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(54) Title: DATA CHECK METHOD, DEVICE AND NETWORK CARD

(54) 发明名称: 数据校验的方法、装置以及网卡



410 ACQUIRE FIRST DATA FROM A SECOND MEMORY OF A SECOND APPARATUS, WHEREIN THE FIRST DATA IS AT LEAST A PART OF DATA IN TARGET DATA TO BE TRANSMITTED
420 BUFFER THE FIRST DATA IN A BUFFER OF A FIRST RNIC
450 THE FIRST RNIC PERFORMS DATA PROCESSING ON THE FIRST DATA TO OBTAIN PROCESSED FIRST DATA, WHEREIN THE PROCESSED FIRST DATA IS DIFFERENT FROM THE FIRST DATA
460 THE FIRST RNIC STORES THE PROCESSED FIRST DATA TO A FIRST MEMORY, WHEREIN A MEMORY ADDRESS STORING THE PROCESSED FIRST DATA IN THE FIRST MEMORY IS THE SAME AS A MEMORY ADDRESS STORING THE FIRST DATA IN THE FIRST MEMORY
430 AFTER THE FIRST RNIC STORES THE TARGET DATA TO THE FIRST MEMORY OF THE FIRST APPARATUS, THE FIRST RNIC READS SECOND DATA IN THE TARGET DATA FROM THE FIRST MEMORY, WHEREIN THE SECOND DATA IS DATA HAVING THE SAME POSITION AS THE FIRST DATA OF THE TARGET DATA
440 IF THE FIRST RNIC DETERMINES THAT THE FIRST DATA IS IDENTICAL TO THE SECOND DATA, THE FIRST RNIC SENDS AN ACK TO THE SECOND RNIC
AA FIRST RNIC OF FIRST APPARATUS
BB SECOND RNIC OF SECOND APPARATUS

(57) Abstract: The present application provides a data check method, device and a network card. A first RNIC determines whether first data is identical to second data, checks data stored in a first memory, determines that first data in target data to be transmitted is correctly stored in the first memory, and notifies a second RNIC by use of an acknowledgement character (ACK). Compared to the prior art in which a first RNIC directly sends an ACK to a second RNIC after receiving a last RDMA packet, the present application improves reliability of storage of target data to a first apparatus by a second apparatus, basically by improving the reliability of the ACK

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

sent from the first apparatus to the second apparatus. The ACK sent from the first apparatus to the second apparatus can be further used to indicate that the first data in the target data is correctly stored in the first memory and helps improve the reliability of direct memory access (DMA) of RDMA for an upper-layer service application.

(57) 摘要: 本申请提供了一种数据校验的方法、装置和网卡, 第一RNIC通过确定第一数据和第二数据是否相同, 对存储在第一内存中的数据进行了校验, 确定待传输的目标数据中第一数据被正确存储到第一内存中, 并通过确认字符ACK通知第二RNIC, 相对于现有技术中, 第一RNIC在接收到最后一个RDMA报文之后直接向第二RNIC发送ACK的方式而言, 有利于提高第二设备对目标数据在第一设备中的存储结果的可信度, 也就是说增强了第一设备向第二设备发送的ACK的可信度。上述第一设备向第二设备发送的ACK还可以用于指示目标数据中的第一数据被正确存储在第一内存中, 有利于提高上层业务应用对RDMA的直接数据存取(Direct Memory Access, DMA)结果的信任度。

数据校验的方法、装置以及网卡

技术领域

5 本申请涉及存储领域，并且更具体地，涉及数据校验的方法、装置以及网卡。

背景技术

10 随着在网络上传输的 NVMe (Non-Volatile Memory express over fabric, NVMf) 技术的发展，远程直接数据存取 (Remote Direct Memory Access, RDMA) 在存储领域的低时延技术中将发挥越来越重要的作用。RDMA 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数据传输技术。RDMA 通过网络把资料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任何影响。

15 基于 RDMA 的上层协议 (Upper Level Protocol, ULP) 业务通常会将访问存储空间的输入输出 IO 业务分解成 RDMA Send、RDMA Read 操作和 RDMA Write 操作，其中 RDMA Send 用于传输 IO 控制命令，RDMA Read 和 RDMA Write 用于传输 IO 数据 (即，待传输的目标数据)。

20 在第一设备与第二设备之间传输目标数据的过程中，第二设备中可以实现 RDMA 功能的网络适配器 (RDMA enable network interface controller, RNIC) 会将需要传输的目标数据封装成至少一个 RDMA 报文，并将上述至少一个 RDMA 报文通过 RDMA 消息发送至第一设备的 RNIC，第一设备的 RNIC 在接收到上述至少一个 RDMA 报文中的最后一个 RDMA 报文后，第一设备的 RNIC 直接向第二设备的 RNIC 发送 ACK。

25 然而，第一设备的 RNIC 直接向第二设备的 RNIC 发送的确认字符 (Acknowledgement, ACK) 仅能表示第一设备接收到最后一个 RDMA 报文，而第二设备无法通过第一设备发送的 ACK 确定 RDMA 报文中的目标数据是否被存储到了第一设备的内存中。

发明内容

本申请提供一种数据校验的方法、装置以及网卡，在一定程度上可以验证目标数据中的至少部分是否被存入第一设备的内存中，有利于提高第二设备对目标数据在第一设备中的存储结果的可信度。

30 第一方面，提供一种数据校验的方法，包括：第一设备的第一远程直接数据存取网络接口卡 RNIC 通过第二设备的第二 RNIC，从所述第二设备的第二内存中获取第一数据，所述第一数据为待传输的目标数据中的至少部分数据；所述第一 RNIC 将所述第一数据缓存在所述第一 RNIC 的缓存中；在所述第一 RNIC 将所述目标数据存储到所述第一设备的第一内存中之后，所述第一 RNIC 从所述第一内存中读取所述目标数据中的第二数据，所述第二数据和所述第一数据为所述目标数据中位置相同的数据；若所述第一 RNIC 35 确定所述第一数据和所述第二数据相同，所述第一 RNIC 向所述第二 RNIC 发送 ACK。

本申请实施例中，第一 RNIC 通过确定第一数据和第二数据是否相同，对存储在所述第一内存中的数据进行校验，确定待传输的目标数据中第一数据被正确存储到第一内存中，并通过 ACK 通知第二 RNIC，相对于现有技术中，第一 RNIC 在接收到最后一个 RDMA

报文之后直接向第二 RNIC 发送 ACK 的方式而言,有利于提高第二设备对目标数据在第一设备中的存储结果的可信度,也就是说增强了第一设备向第二设备发送的 ACK 的可信度。上述第一设备向第二设备发送的 ACK 还可以用于指示目标数据中的第一数据被正确存储在第二内存中,有利于提高上层业务应用对 RDMA 的直接数据存取 (Direct Memory Access, DMA) 结果的信任度。

进一步地,当上述第一数据为部分目标数据时,也就是说,仅验证一部分目标数据是否被存入第二内存中,对第二 RNIC 中缓存的容量的要求较低,在一定程度上可以节约缓存的成本。

若上述第一数据是目标数据的全部数据时,相当于对目标数据中的全部数据进行了验证,有利于确定目标数据是否全部被正确存储在第二内存中,相对于只验证部分目标数据的方案而言,进一步提高了 ACK 的可信度。

可选地,所述 ACK 用于指示所述第一数据和所述第二数据相同,也就是说,上述第一数据被存储在第二内存中。

可选地,上述第一数据可以是以目标数据在第二设备的第二内存中对应的内存地址的末地址为起点,向所述目标数据在第二设备的第二内存中对应的内存地址的首地址的方向,读取预设的内存地址长度后得到的数据,相应地,上述第二数据也可以是以目标数据在第一设备的第一内存中对应的内存地址的末地址为起点,向所述目标数据在第一设备的第一内存中对应的内存地址的首地址的方向,读取预设的内存地址长度后得到的数据。

通过以存储目标数据的内存地址的末地址到存储目标数据的内存地址的首地址的方向,读取预设内存地址长度的数据作为被验证的数据 (第二数据),第一 RNIC 无需记录存储目标数据的内存地址的首地址,有利于降低第一 RNIC 实现方案的复杂度。

结合第一方面,在第一方面的一种可能的实现方式中,在所述第一 RNIC 将所述目标数据存储到所述第二内存的过程中,且在所述第一 RNIC 将所述目标数据中的第一数据存储到所述第二内存之前,所述方法还包括:所述第一 RNIC 对所述第一数据进行数据处理得到处理后的第一数据,所述处理后的第一数据与所述第一数据不同;所述第一 RNIC 将所述处理后的第一数据存储到所述第二内存中,所述第二内存中存储所述处理后的第一数据的内存地址与所述第二内存中存储所述第一数据的内存地址相同。

因此,为了避免在第一数据存储到第二内存之前,第二内存有可能存储了与第一数据相同的数据的特殊情况,本申请实施例通过对第一数据进行数据处理,并将处理后的数据存入第二内存中,以确保在第一数据存储到第二内存之前,第二内存中存储的数据一定与第一数据不同,此时再结合数据顺序存储的存储机制,可以确定第二内存中内存地址位于第一数据的内存地址之前的数据被存储到第二内存中,有利于进一步提高第二设备对目标数据在第一设备中的存储结果的可信度。

结合第一方面,在第一方面的一种可能的实现方式中,所述目标数据承载于至少一个远程直接数据存取 RDMA 报文中,所述至少一个 RDMA 报文包括第一 RDMA 报文,所述第一 RDMA 报文为传输所述至少一个 RDMA 报文的过程中的最后一个 RDMA 报文,所述第一设备的第一远程直接数据存取网络接口卡 RNIC 通过第二设备的第二 RNIC,从所述第二设备的第二内存中获取第一数据,包括:所述第一 RNIC 接收所述第二 RNIC 发送的所述第一 RDMA 报文,所述第一 RDMA 报文中承载所述第一数据。

因此,本申请实施例基于上述 RDMA 报文的顺序存储机制,通过对最后一个 RDMA 报文中的数据进行验证,可以进一步确定最后一个 RDMA 报文以及在最后一个 RDMA 报文之前传输的 RDMA 报文是否被存入第一内存中,在一定程度上,简化了数据校验的复杂度,还有利于降低实施本申请实施例的方案对于缓存容量的需求。

5 结合第一方面,在第一方面的一种可能的实现方式中,所述目标数据承载于多个 RDMA 报文中,所述多个 RDMA 报文包括第二 RDMA 报文和至少一个第三 RDMA 报文,所述第二 RDMA 报文为传输所述多个 RDMA 报文的过程中的最后一个 RDMA 报文,所述至少一个第三 RDMA 报文为传输所述多个 RDMA 报文的过程中传输所述最后一个 RDMA 报文之前的报文,所述第一设备的第一远程直接数据存取网络接口卡 RNIC 通过
10 第二设备的第二 RNIC,从所述第二设备的第二内存中获取第一数据,包括:所述第一 RNIC 接收所述第二 RNIC 发送的所述第二 RDMA 报文和所述至少一个第三 RDMA 报文,所述第二 RDMA 报文中承载的数据和所述至少一个第三 RDMA 报文中承载的数据组成所述第一数据。

15 结合第一方面,在第一方面的一种可能的实现方式中,所述至少一个第三 RDMA 报文为一个 RDMA 报文,所述第三 RDMA 报文为传输所述多个 RDMA 报文的过程中传输所述最后一个 RDMA 报文的前一个 RDMA 报文。

本申请实施例中,通过对最后一个 RDMA 报文以及最后一个 RDMA 报文的前一个 RDMA 报文中的数据进行验证,相对于随机对目标数据中的任意一段数据进行验证的方案而言,有利于降低第一 RNIC 进行数据校验的复杂度。

20 第二方面,提供了一种数据校验的装置,所述数据校验的装置包括用于执行第一方面或第一方面任一种可能实现方式中的各个模块。

第三方面,提供了一种网卡,所述网卡设置在第一设备中,所述网卡包括处理器、存储器以及收发器,所述存储器存储有程序指令,所述处理器运行所述程序指令以执行以下步骤:控制所述收发器通过第二设备的第二网卡,从所述第二设备的第二内存中获取第一数据,所述第一数据为待传输的目标数据中的至少部分数据;将所述第一数据缓存在所述网卡的缓存中;将所述目标数据存储到所述第一设备的第一内存中之后,从所述
25 第一内存中读取所述目标数据中的第二数据,所述第二数据和所述第一数据为所述目标数据中位置相同的数据;若确定所述第一数据和所述第二数据相同,控制所述收发器向所述第二 RNIC 发送确认字符 ACK。

30 第四方面,提供了一种芯片,包括收发器、处理器和存储器。该处理器用于控制收发器收发信号,该存储器用于存储计算机程序,该处理器用于从存储器中调用并运行该计算机程序,使得该数据校验的装置执行第一方面中的方法。

第五方面,提供了一种计算机可读存储介质,所述计算机可读存储介质中存储有指令,当其在计算机上运行时,使得计算机执行上述各方面所述的方法。

35 第六方面,提供了一种包含指令的计算机程序产品,当其在计算机上运行时,使得计算机执行上述各方面所述的方法。

本申请实施例中,第一 RNIC 通过确定第一数据和第二数据是否相同,对存储在
40 第一内存中的数据进行校验,确定待传输的目标数据中第一数据被正确存储到第一内存中,并通过 ACK 通知第二 RNIC,有利于提高第二设备对目标数据在第一设备中的存储结果的可信度,也就是说增强了第一设备向第二设备发送的 ACK 的可信度。

附图说明

图 1 是本申请实施例的传输数据的装置的示意性框图；

图 2 是基于 RDMA 的 IO 读过程的方法的示意性流程图；

5 图 3 是基于 RDMA 的 IO 写过程的方法的示意性流程图；

图 4 是本申请实施例的数据校验的方法的示意性流程图；

图 5 是本申请实施例中基于 IO 读过程的数据校验的方法的示意性流程图；

图 6 是本申请实施例中基于 IO 写过程的数据校验的方法的示意性流程图；

图 7 是本申请实施例的数据校验的装置的示意性框图；

10 图 8 是本申请另一实施例的网卡的示意性框图。

具体实施方式

下面将结合附图，对本申请中的技术方案进行描述。

15 为了便于理解本申请实施例的方法，先结合图 1 介绍下本申请实施例的传输数据的装置的示意性结构图。

图 1 是本申请实施例的传输数据的装置的示意性框图。图 1 所示的传输数据的装置 100 可以包括：内存 110、中央处理器 120、芯片组 130 和 RNIC140。其中，内存 110、处理器 120、芯片组 130 和 RNIC140 通过内部连接通路相连。

内存 110，用于存储数据和指令。

20 处理器 120，用于解释内存 110 中存储的指令以及处理内存 110 中存储的数据，是上述数据校验的装置的控制核心和计算核心。

芯片组 (Chipset) 130，用于与处理器 120 协同工作，解释内存 110 存储的指令或处理内存 110 中的数据。

RNIC 140，用于将数据校验装置接入网络，与接入该网络的其他设备进行通信。

25 下文基于图 1 所示的装置结合图 2 和图 3 详细描述传统的数据传输方法。应理解，图 1 所示的装置可以是图 2 和图 3 中所示的发起数据 IO 过程的发起端，图 1 所示的装置还可以是图 2 和图 3 所示的被发起端写入数据或读取数据的目标端。

需要说明的是，图 2 和图 3 中的发起端和目标端可以是存储系统中的服务器、存储控制器、或存储设备等，本申请对上述发起端和目标端的具体形式不做限定。

30 图 2 是基于 RDMA 的 IO 读过程的方法的示意性流程图。图 2 所示的方法包括步骤 210-280。

210，发起端 (Initiate, INI) 中基于 RDMA 的 ULP 业务发起 IO 读请求 (IO Read)。

220，发起端根据 IO 读请求，生成 RDMA 发送 (RDMA Send) 命令。

35 具体地，上述 RDMA Send 命令包括 IO 读请求，以及发起端的第一目标缓存的描述信息，其中第一目标缓存用于存储通过 IO 读请求读取的目标数据。上述第一目标缓存的描述信息可以是第一目标缓存的主键 (key)，第一目标缓存空间的虚拟地址 (Virtual Address, VA)，第一目标缓存空间的长度 (Length) 等信息。

230，发起端向目标端 (Target, TGT) 发送 RDMA Send 命令。

40 具体地，目标端从 RDMA Send 命令中解析 IO 读请求，并提取第一目标缓存的描述信息。

240, 目标端根据 IO 读请求以及第一目标缓存的描述信息, 从目标端的缓存中确定第二目标缓存, 第二目标缓存用于缓存通过 IO 读请求从目标端的内存中读取的数据。

250, 目标端根据 IO 读请求, 将目标数据从目标端的内存中读到目标端的第二目标缓存中。

5 例如, 目标端可以根据 IO 读请求中携带的目标数据的存储地址, 确定目标数据在目标端的内存中的存储地址, 从而将该内存中的存储地址对应的目标数据从目标端的内存中读到目标端的第二目标缓存中。

260, 目标端的 RNIC 将第二目标缓存中的目标数据封装成至少一个 RDMA 报文, 并通过 RDMA Write 命令将 RDMA 报文发送至发起端的 RNIC。

10 270, 发起端的 RNIC 接收到至少一个 RDMA 报文中的最后一个 RDMA 报文后, 发起端的 RNIC 向目标端的 RNIC 发送 ACK。

280, 目标端的 RNIC 通过 RDMA Send 向发起端的 RNIC 发送 IO 响应(response)。

具体地, 发起端的 RNIC 接收到 IO 响应后, 发起端和目标端之间的数据传输的流程结束。

15 图 3 是基于 RDMA 的 IO 写过程的方法的示意性流程图。图 3 所示的方法包括步骤 310-380。

310, 发起端(Initiate, INI)中基于 RDMA 的 ULP 业务发起 IO 写请求(IO Write)。

320, 发起端根据 IO 写请求, 生成 RDMA Send 命令。

20 具体地, 上述 RDMA Send 命令包括 IO 写请求, 以及发起端的第一目标缓存的描述信息, 其中第一目标缓存用于存储准备通过 IO 写请求向目标端写入的目标数据。上述第一目标缓存的描述信息可以是第一目标缓存的主键(key), 第一目标缓存的虚拟地址(Virtual Address, VA), 第一目标缓存的长度(Length)等信息。

330, 发起端向目标端发送 RDMA Send 命令。

25 具体地, 目标端从 RDMA Send 命令中解析 IO 写请求, 并提取发起端中第一目标缓存的描述信息。

340, 目标端根据接收到的 IO 写请求以及第一目标缓存的描述信息, 从目标端的缓存中确定第二目标缓存, 该第二目标缓存用于缓存发起端通过 IO 写请求写入目标端的目标数据。

350, 目标端向发起端发送 RDMA 写请求(Read Request)。

30 360, 发起端从发起端的内存中将准备通过 IO 写请求传输的目标数据读到第一目标缓存中。

370, 发起端的 RNIC 将第一目标缓存中的目标数据封装成至少一个 RDMA 报文, 并通过 RDMA 读响应(Read Response)命令将 RDMA 报文发送至目标端的 RNIC。

35 380, 目标端的 RNIC 接收到至少一个 RDMA 报文中的最后一个 RDMA 报文后, 目标端的 RNIC 向发起端的 RNIC 发送 ACK。

390, 目标端的 RNIC 通过 RDMA Send 向发起端的 RNIC 发送 IO 响应(response)。

具体地, 发起端的 RNIC 接收到 IO 响应后, 发起端和目标端之间的数据传输的流程结束。

40 为了便于描述, 下文中将接收 RDMA 报文的设备称为“第一设备”, 将发送 RDMA 报文的设备称为“第二设备”, 例如, 在 IO 写过程中第一设备可以是上文中的目标端, 第

二设备可以是上文中的发起端；在 IO 读过程中第一设备可以是上文中的发起端，第二设备可以是上文中的目标端。

从图 2 和图 3 所示的基于 RDMA 的 IO 请求（写请求或读请求）进行数据传输的过程中，第一设备的 RNIC 在接收到第二设备的 RNIC 发送的承载目标数据的最后一个 RDMA 报文后，会直接向第二设备的 RNIC 发送 ACK。然而，第二设备通过 ACK 仅能确定第一设备的 RNIC 接收到承载目标数据的最后一个 RDMA 报文，无法确定第一设备的 RNIC 是否将目标数据存储到第一设备的内存中。

为了解决上述问题，下文结合图 4 至图 8，详细描述本申请实施例的数据校验的方法和装置。

图 4 是本申请实施例的数据校验的方法的示意性流程图。图 4 所示的方法包括：

410，第一设备的第一远程直接数据存取网络接口卡 RNIC，通过第二设备的第二 RNIC，从所述第二设备的第二内存中获取第一数据，所述第一数据为待传输的目标数据中的至少部分数据。

具体地，上述 RNIC 可以是能够实现 RDMA 技术和底层协议（Lower Layer Protocol, LLP）的网络适配器（Network Interface Card, NIC）。

上述待传输的目标数据可以是第一设备的第一 RNIC 需要直接从第二设备的内存中读取的数据。

上述第一数据为待传输的目标数据中的至少部分数据，可以指第一数据是待传输的目标数据中的全部数据；还可以指第一数据是待传输的目标数据中的任意一段数据，或者第一数据是待传输的目标数据中的任意 1 字节数据。

可选地，上述第一数据可以是以目标数据在第二设备的第二内存中对应的内存地址的末地址为起点，向所述目标数据在第二设备的第二内存中对应的内存地址的首地址的方向，读取预设的内存地址长度后得到的数据。

应理解，上述内存地址可以是逻辑地址，也就是说，存储目标数据的内存地址（逻辑地址）可以是连续的，但是存储目标数据的物理地址可以是离散的也可以是连续的。

需要说明的是，上述内存可以按照字节编制，也就是说，内存中的每个地址可以表示一个字节，则预设的内存地址长度可以以字节为单位；内存中每个字节可以使用 8 比特位表示，则上述预存的内存地址长度还可以是比特位为单位，本申请实施例对于上述预设的内存地址长度的具体体现形式不做限定。

例如，目标数据在第二设备的第二内存中的内存地址的末地址为 N，且预设内存地址长度为 4 字节，则第一数据可以是存储在内存地址为 N、N-1、N-2 和 N-4 上的数据。

可选地，作为一个实施例，所述目标数据承载于至少一个远程直接数据存取 RDMA 报文中，所述至少一个 RDMA 报文包括第一 RDMA 报文，所述第一 RDMA 报文为传输所述至少一个 RDMA 报文的过程中的最后一个 RDMA 报文，步骤 410 包括：所述第一 RNIC 接收所述第二 RNIC 发送的所述第一 RDMA 报文，所述第一 RDMA 报文中承载所述第一数据。

具体地，若待传输的目标数据承载在一个 RDMA 报文中，承载该目标数据的 RDMA 报文可以称为 Only 报文，则上述第一数据可以为 Only 报文中的全部数据；若承载目标数据的 RDMA 报文可以称为 Only 报文，则上述第一数据还可以为 Only 报文中的部分数据。

若待传输的目标数据承载在多个 RDMA 报文中,则承载第一数据的报文可以是传输上述多个 RDMA 报文的过程中最后传输的 RDMA 报文,即最后一个 RDMA 报文。上述第一数据可以是最后一个 RDMA 报文中携带的部分数据,上述第一数据还可以是最后一个 RDMA 报文中携带的全部数据。

5 举例而言,发起端的 RNIC 可通过检测 RDMA 报文的是否携带有 last 标识来判断接收到 RDMA 报文是否是最后一个 RDMA 报文。

还应理解,上述第一数据的数据长度可以是预设的,例如,预设第一数据的数据长度为 1 字节,则可以从上述第一 RDMA 报文中提取任意 1 字节作为第一数据。上述第一数据的数据长度还可以是随上述第一 RDMA 报文中携带的数据长度变化而变化的,也就是说,当第一 RDMA 报文中携带了 4 字节的数据,则第一数据的数据长度可以是 4 字节。

10 需要说明的是,上述 RDMA 报文可以指基于 RDMA 技术传输的数据报文,又称 RDMA 数据包(RDMA Packet)。

在传输多个 RDMA 报文的过程中,多个 RDMA 报文是可以按照预设的顺序(例如, RDMA 报文的生成数据)依次被存入第一内存中的,也就是说,在第 N-1 个 RDMA 报文被存储第一内存中之前,第 N 个 RDMA 报文是不会被存入第一内存中的,其中, N 为自然数,且第 N 个 RDMA 报文是第 N-1 个 RDMA 报文的下一个 RDMA 报文。

15 因此,本申请实施例基于上述 RDMA 报文的顺序存储机制,通过对最后一个 RDMA 报文中的数据进行验证,可以进一步确定最后一个 RDMA 报文以及在最后一个 RDMA 报文之前传输的 RDMA 报文是否被存入第一内存中,在一定程度上,简化了数据校验的复杂的,还有利于降低本申请实施例的方案对于缓存容量的需求。

20 可选地,作为一个实施例,所述目标数据承载于多个 RDMA 报文中,所述多个 RDMA 报文包括第二 RDMA 报文和至少一个第三 RDMA 报文,所述第二 RDMA 报文为传输所述多个 RDMA 报文的过程中的最后一个 RDMA 报文,所述至少一个第三 RDMA 报文为传输所述多个 RDMA 报文的过程中传输所述最后一个 RDMA 报文之前的报文,步骤 410 包括:所述第一 RNIC 接收所述第二 RNIC 发送的所述第二 RDMA 报文和所述至少一个第三 RDMA 报文,所述第二 RDMA 报文中承载的数据和所述至少一个第三 RDMA 报文中承载的数据组成所述第一数据。

25 具体的,上述第二 RDMA 报文中承载的数据和所述至少一个第三 RDMA 报文中承载的数据组成所述第一数据,可以指第一数据是由第二 RDMA 报文中的全部数据和至少一个第三 RDMA 报文中的至少部分数据组成的;上述第一数据还可以是由第二 RDMA 报文中的部分数据和至少一个第三 RDMA 报文中的至少数据组成的。

需要说明的是,在上文中提到上述第一数据的数据长度可以是预设的,当第二 RDMA 报文中携带的数据的数据长度不足上述预设的数据长度时,可以从至少一个第三 RDMA 报文中提取至少部分数据与第二 RDMA 报文中携带的数据一起补足预设的数据长度。

30 例如,预设的第一数据的长度为 4 字节,第二 RDMA 报文中携带的数据的数据长度为 1 字节,此时还需要从至少一个第三 RDMA 报文中提取 3 字节的数据与第二 RDMA 报文中的数据一起组成第一数据,满足预设的第一数据的长度。

35 还应理解,从上述至少一个第三 RDMA 报文中提取至少部分数据与所述第二 RDMA 报文中的数据组成第一数据时,上述至少一个第三 RDMA 中的部分数据可以随机进行选择,例如,可以选择在传输承载目标数据的多个 RDMA 报文的过程中,最开始传输的第

一 RDMA 报文中的至少部分数据。

可选地，作为一个实施例，所述至少一个第三 RDMA 报文为一个 RDMA 报文，所述第三 RDMA 报文为传输所述多个 RDMA 报文的过程中传输所述最后一个 RDMA 报文的前一个 RDMA 报文。

5 可选地，承载上述目标数据的多个 RDMA 报文的传输顺序与所述多个 RDMA 报文中携带的数据在第二内存中的内存地址的递增的顺序相同。换句话说，传输多个 RDMA 报文中的第 N-1 个 RDMA 报文和第 N 个 RDMA 报文时，第 N-1 个 RDMA 报文中携带的数据的内存地址的取值的最大值小于第 N 个 RDMA 报文中携带的数据的内存地址的取值的最小值，且所述多个 RDMA 报文中承载的目标数据在第二内存中的内存地址是连续的，所述多个 RDMA 报文中的每个 RDMA 报文中的数据在第二内存中的内存地址也是连续的。

例如，承载待传输的目标数据的 RDMA 报文为 3 个 RDMA 报文，包括 RDMA 报文 1、RDMA 报文 2 和 RDMA 报文 3，传输上述 3 个 RDMA 报文的顺序也是按照从 RDMA 报文 1 到 RDMA3 的顺序依次传输的，也就是说，RDMA 报文 3 是传输 3 个 RDMA 报
15 文中的最后一个 RDMA 报文。上述每个 RDMA 报文中携带 2 字节的数据，且预设的第一数据的数据长度为 4 字节，则 RDMA 报文 3 中携带的数据的数据长度不满足预设的第一数据的数据长度，此时，可以从 RDMA 报文 2 中提取 2 字节的数据与 RDMA 报文 3 中的数据一起组成第一数据。

420，所述第一 RNIC 将所述第一数据缓存在所述第一 RNIC 的缓存中。

20 430，在所述第一 RNIC 将所述目标数据存储到所述第一内存中之后，所述第一 RNIC 从所述第一内存中读取第二数据，所述第二数据和所述第一数据为所述目标数据中位置相同的数据。

具体地，上述第二数据和第二数据为目标数据中位置相同的数据可以理解为第二数据在所述第一内存中的内存地址在第一内存中存储所述目标数据的内存地址的范围内的
25 相对位置，与所述第一数据在所述第二内存中的内存地址在第二内存中存储所述目标数据的内存地址的范围内的相对位置相同；或者第二数据在第一内存中的内存地址相对于目标数据在第一内存中的内存地址的首地址的偏移量，与第一数据在第二内存中的内存地址相对于目标数据在第二内存中的内存地址的首地址的偏移量相等；或者，第二数据在第一内存中的内存地址相对于目标数据在第一内存中的内存地址的末地址的偏移量，
30 与第一数据在第二内存中的内存地址相对于目标数据在第二内存中的内存地址的末地址的偏移量相同，换句话说，第二数据和第一数据是原本应该相同的数据。

440，若所述第一 RNIC 确定所述第一数据和所述第二数据相同，所述第一 RNIC 向所述第二 RNIC 发送 ACK。

本申请实施例中，第一 RNIC 通过确定第一数据和第二数据是否相同，对存储在
35 第一内存中的数据进行校验，确定待传输的目标数据中的至少部分数据被正确存储到第一内存中，并通过 ACK 通知第二 RNIC，相对于现有技术中，第一 RNIC 在接收到最后一个 RDMA 报文之后直接向第二 RNIC 发送 ACK 的方式而言，增强了 ACK 的可信度，即该 ACK 可以用于指示目标数据中的至少部分数据被正确存储在第一内存中，有利于提高上层业务应用对 RDMA 的直接数据存取 (Direct Memory Access, DMA) 结果的信任度。

40 进一步地，当上述第一数据为部分目标数据时，也就是说，仅验证一部分目标数据

是否被存入第一内存中，对第一 RNIC 中缓存的容量的要求较低，在一定程度上可以节约缓存的成本。

若上述第一数据是目标数据的全部数据时，相当于对目标数据中的全部数据进行了验证，有利于确定目标数据是否全部被正确存储第一内存中，相对于只验证部分目标数据的方案而言，进一步提高了 ACK 的可信度。

可选地，所述 ACK 用于指示所述第一数据和所述第二数据相同，也就是说，上述第一数据被存储在第一内存中。

需要说明的是，上述第一 RNIC 向所述第二 RNIC 发送 ACK 后，可以理解为数据传输过程结束。

结合图 1 和图 2 可以看出，图 1 或图 2 示出的流程中，数据传输的结束点在步骤 270 或步骤 380。而本申请实施例的方案中，数据传输过程的结束点可停留在步骤 260，或者步骤 370，可以节省第一 RNIC 和第二 RNIC 之间的信令交互。

可选地，作为一个实施例，在步骤 430 之前，所述方法还包括步骤 450 和步骤 460。

450，所述第一 RNIC 对所述第一数据进行数据处理得到处理后的第一数据，所述处理后的第一数据与所述第一数据不同。

具体地，在所述第一 RNIC 将所述目标数据存储到所述第一内存的过程中，且在所述第一 RNIC 将所述目标数据中的第一数据存储到所述第一内存中之前，第一 RNIC 将第一数据进行数据处理得到处理后的第一数据。

需要说明的是，上述对第一数据进行处理可以指对第一数据进行取反计算，使得处理后的第一数据与第一数据不同；或者当第一数据包含多个字节且每个字节中的数据内容不同时，可以通过打乱第一数据中的多个字节的顺序，使得处理后的第一数据与第一数据的内容不同，本申请实施例对于数据处理的具体方式不做具体限定。

460，所述第一 RNIC 将所述处理后的第一数据存储到所述第一内存中，所述第一内存中存储所述处理后的第一数据的内存地址与所述第一内存中存储所述第一数据的内存地址相同。

需要说明的是，目标数据在第二内存中可以存储在一段连续的内存地址对应的内存空间中，第一 RNIC 在向第一内存中存储目标数据时，可以按照目标数据在第二内存中的内存地址顺序将目标数据存入第一内存中，假设目标数据在第二内存中的内存地址包括第一段内存地址和第二段内存地址，且第一段内存地址与第二段内存地址连续，第一段内存地址的偏移量小于第二段内存地址的偏移量，则在目标数据存储到第一内存的过程中，如果第一段内存地址中的数据未存到第一内存中，则第二段内存地址中的数据也不会继续往第一内存中存储。因此，基于上述数据顺序的存储机制，如果确定第二段存储地址中的数据被存入第一内存中，则可以认为第一段存储地址中的数据被存入第一内存中。

若上述的第二段存储地址中的数据为上述第一数据时，在上述第一数据在向第一内存中存储之前，第一内存中可能已经存储有数据，并且在特殊情况下，第一内存中原本存储的数据可能和第一数据的内容相同，此时，无论第一数据是否存储到第一内存中，还是可以得到第一 RNIC 确定缓存中的第一数据和第一内存中读取第二数据相同的结果，在上述这种情况下，其实无法通过上述数据顺序存储的存储机制确定目标数据中存储在

因此，为了避免在第一数据存储到第一内存中之前，第一内存有可能存储了与第一数据相同的数据的特殊情况，本申请实施例通过对第一数据进行数据处理，并将处理后的数据存入第一内存中，以确保在第一数据存储到第一内存中之前，第一内存中存储的数据一定与第一数据不同，此时再结合数据顺序存储的存储机制，可以确定第二内存中内存地址位于第一数据的内存地址之前的数据被存储到第一内存中，以进一步提高上层业务应用对 RDMA 的直接数据存取（Direct Memory Access, DMA）结果的信任度。

为了便于理解，下文结合图 5 和图 6 分别以 IO 读过程和 IO 写过程描述本申请实施例的数据校验方法。需要说明的是，下文中 IO 读过程和 IO 写过程的描述侧重点为与本申请实施例相关的部分，其余可以参见上文中图 2 和图 3 所示的 IO 读过程和 IO 写过程的描述，为了简洁，在此不再赘述。

图 5 是本申请实施例中基于 IO 读过程的数据校验的方法的示意性流程图。图 5 所示的方法包括 510-560。

510，目标端的第一 RNIC，通过发起端的第二 RNIC，从目标端的第二内存中获取第一数据，第一数据为待传输的目标数据中的至少部分数据。

需要说明的是，上述目标端可以是发起 IO 读过程的设备，上述目标端可以是存储需要通过 IO 读过程读取的数据。

520，目标端的第一 RNIC 将第一数据缓存在第一 RNIC 的缓存中。

530，目标端的第一 RNIC 对第一数据进行数据处理，得到处理后的第一数据，处理后的第一数据和第一数据不同。

540，目标端第一 RNIC 将处理后的第一数据存储在第一内存中，第一内存中存储所述处理后的第一数据的内存地址与第一内存中存储第一数据的内存地址相同。

550，在目标端的第一 RNIC 将目标数据存储到目标端的第一内存中之后，第一 RNIC 从第一内存中读取第二数据，第二数据和第一数据为目标数据中位置相同的数据。

560，若目标端的第一 RNIC 确定第一数据和第二数据相同，目标端的第一 RNIC 向发起端的第二 RNIC 发送 ACK。

图 6 是本申请实施例中基于 IO 写过程的数据校验的方法的示意性流程图。图 6 所示的方法包括 610-660。

610，发起端的第一 RNIC，通过目标端的第二 RNIC，从目标端的第二内存中获取第一数据，第一数据为待传输的目标数据中的至少部分数据。

需要说明的是，上述发起端可以是发起 IO 读过程的设备，上述目标端可以是存储需要通过 IO 读过程读取的数据。

620，发起端的第一 RNIC 将第一数据缓存在第一 RNIC 的缓存中。

630，发起端的第一 RNIC 对第一数据进行数据处理，得到处理后的第一数据，处理后的第一数据和第一数据不同。

640，发起端第一 RNIC 将处理后的第一数据存储在第一内存中，第一内存中存储所述处理后的第一数据的内存地址与第一内存中存储第一数据的内存地址相同。

650，在发起端的第一 RNIC 将目标数据存储到发起端的第一内存中之后，第一 RNIC 从第一内存中读取第二数据，第二数据和第一数据为目标数据中位置相同的数据。

660，若发起端的第一 RNIC 确定第一数据和第二数据相同，发起端的第一 RNIC 向目标端的第二 RNIC 发送 ACK。

上文结合图 1 至图 6 详细地描述了本申请实施例的数据校验的方法，下文结合图 7 和图 8 详细地描述本申请实施例的数据校验的装置。

图 7 是本申请实施例的数据校验的装置的示意性框图。图 7 所示的数据校验的装置 700 包括获取单元 710、处理单元 720 和发送单元 730。

5 获取单元，用于通过第二设备的第二 RNIC 从所述第二设备的第二内存中获取第一数据，所述第一数据为待传输的目标数据中的至少部分数据；

处理单元，用于将所述第一数据缓存在第一设备的第一 RNIC 的缓存中；

所述处理单元，还用于从第一设备的内存中读取第二数据，所述第二数据在所述第一内存中的内存地址在第一内存中存储所述目标数据的内存地址的范围内的相对位置，
10 与所述第一数据在所述第二内存中的内存地址在第二内存中存储所述目标数据的内存地址的范围内的相对位置相同；

若所述第一 RNIC 确定所述第一数据和所述第二数据相同，发送单元，用于向所述第二 RNIC 发送 ACK。

可选地，作为一个实施例，所述处理单元，还用于将所述第一数据进行处理得到处理后的第一数据，所述处理后的第一数据与所述第一数据不同；所述处理单元，还用于将所述处理后的第一数据存储在第一内存中，所述第一内存中存储所述处理后的第一数据的内存地址与所述第一内存中存储所述第一数据的内存地址相同。
15

可选地，作为一个实施例，承载所述目标数据的报文包括至少一个远程直接数据存取 RDMA 报文，所述至少一个 RDMA 报文包括第一 RDMA 报文，所述第一 RDMA 报文为传输所述至少一个 RDMA 报文的过程中的最后一个 RDMA 报文，
20

所述获取单元，具体用于接收所述第二 RNIC 发送的所述第一 RDMA 报文，所述第一 RDMA 报文中承载所述第一数据。

可选地，作为一个实施例，承载所述目标数据的报文包括多个 RDMA 报文，所述多个 RDMA 报文包括第二 RDMA 报文和至少一个第三 RDMA 报文，所述第二 RDMA 报文为传输所述多个 RDMA 报文的过程中的最后一个 RDMA 报文，所述至少一个第三 RDMA 报文为传输所述多个 RDMA 报文的过程中传输所述最后一个 RDMA 报文之前的报文，
25

所述获取单元，具体用于接收所述第二 RNIC 发送的所述第二 RDMA 报文和所述至少一个第三 RDMA 报文，所述第二 RDMA 报文中承载的数据和所述至少一个第三 RDMA 报文中承载的数据组成所述第一数据。
30

可选地，作为一个实施例，所述至少一个第三 RDMA 报文为一个 RDMA 报文，所述第三 RDMA 报文为传输所述多个 RDMA 报文的过程中传输所述最后一个 RDMA 报文的前一个报文。

在可选的实施例中，所述数据校验的装置 700 可以为网卡 800，所述获取单元 710 和所述发送单元 730 可以为网卡中的收发器 870，所述处理单元 720 可以为网卡中的处理器 820，所述网卡还可以包括输入/输出接口 830 和存储器 810，具体如图 8 所示。
35

图 8 是本申请另一实施例的网卡的示意性框图。图 8 所示的网卡 800 可以包括：存储器 810、处理器 820、输入/输出接口 830、收发器 840。其中，存储器 810、处理器 820、输入/输出接口 830 和收发器 840 通过内部连接通路相连，该存储器 810 用于存储程序指令，该处理器 820 用于执行该存储器 820 存储的程序指令，以控制输入/输出接口 830 接
40

收输入的数据和信息，输出操作结果等数据，并控制收发器 840 发送信号。

所述处理器运行所述程序指令以执行上文第一设备的第一 RNIC 所执行的步骤，例如，可控制所述收发器通过第二设备的第二网卡，从所述第二设备的第二内存中获取第一数据，所述第一数据为待传输的目标数据中的至少部分数据；将所述第一数据缓存在所述网卡的缓存中；将所述目标数据存储到所述第一设备的第一内存中之后，从所述第一内存中读取所述目标数据中的第二数据，所述第二数据和所述第一数据为所述目标数据中位置相同的数据；若确定所述第一数据和所述第二数据相同，控制所述收发器向所述第二 RNIC 发送确认字符 ACK。

可选地，上述的缓存可设置在存储器中。

应理解，在本申请实施例中，该处理器 820 可以采用通用的中央处理器（Central Processing Unit, CPU），微处理器，应用专用集成电路（Application Specific Integrated Circuit, ASIC），或者一个或多个集成电路，用于执行相关程序，以实现本申请实施例所提供的技术方案。

还应理解，收发机 840 又称通信接口，使用例如但不限于收发器一类的收发装置，来实现数据校验的装置 800 与其它设备或通信网络之间的通信。

该存储器 810 可以包括只读存储器和随机存取存储器，并向处理器 820 提供指令和数据。处理器 820 的一部分还可以包括非易失性随机存取存储器。例如，处理器 820 还可以存储设备类型的信息。

在实现过程中，上述方法的各步骤可以通过处理器 820 中的硬件的集成逻辑电路或者软件形式的指令完成。结合本申请实施例所公开的数据校验方法可以直接体现为硬件处理器执行完成，或者用处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器 810，处理器 820 读取存储器 810 中的信息，结合其硬件完成上述方法的步骤。为避免重复，这里不再详细描述。

应理解，本申请实施例中，该处理器可以为中央处理单元（central processing unit, CPU），该处理器还可以是其它通用处理器、数字信号处理器（digital signal processor, DSP）、专用集成电路（application specific integrated circuit, ASIC）、现成可编程门阵列（field programmable gate array, FPGA）或者其它可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

在可选的实施例中，上述的数据校验的装置可以是位于存储系统 600 中的 RNIC，该存储系统还包括内存（Memory），芯片组（Chipset）和中央处理器（Central Processing Unit, CPU）。

应理解，在本申请实施例中，“与 A 相应的 B”表示 B 与 A 相关联，根据 A 可以确定 B。但还应理解，根据 A 确定 B 并不意味着仅仅根据 A 确定 B，还可以根据 A 和/或其它信息确定 B。

应理解，本文中术语“和/或”，仅仅是一种描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。另外，本文中字符“/”，一般表示前后关联对象是一种“或”的关系。

应理解，在本申请的各种实施例中，上述各过程的序号的大小并不意味着执行顺序

的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不应对本申请实施例的实施过程构成任何限定。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元5的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示10的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。15当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读20存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（Digital Subscriber Line, DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够读取的任何可用介质或者是包含一个或多个可用介质集成的25的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，数字通用光盘（Digital Video Disc, DVD））或者半导体介质（例如，固态硬盘（Solid State Disk, SSD））等。

以上所述，仅为本发明的具体实施方式，但本发明的保护范围并不局限于此，任何30熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应以所述权利要求的保护范围为准。

权利要求

1、一种数据校验的方法，其特征在于，包括：

第一设备的第一远程直接数据存取网络接口卡 RNIC 通过第二设备的第二 RNIC，从所述第二设备的第二内存中获取第一数据，所述第一数据为待传输的目标数据中的至少部分数据；

所述第一 RNIC 将所述第一数据缓存在所述第一 RNIC 的缓存中；

在所述第一 RNIC 将所述目标数据存储到所述第一设备的第一内存中之后，所述第一 RNIC 从所述第一内存中读取所述目标数据中的第二数据，所述第二数据和所述第一数据为所述目标数据中位置相同的数据；

若所述第一 RNIC 确定所述第一数据和所述第二数据相同，所述第一 RNIC 向所述第二 RNIC 发送确认字符 ACK。

2、如权利要求 1 所述的方法，其特征在于，在所述第一 RNIC 将所述目标数据存储到所述第一内存的过程中，且在所述第一 RNIC 将所述目标数据中的第一数据存储到所述第一内存中之前，所述方法还包括：

所述第一 RNIC 对所述第一数据进行数据处理得到处理后的第一数据，所述处理后的第一数据与所述第一数据不同；

所述第一 RNIC 将所述处理后的第一数据存储在第一内存中，所述第一内存中存储所述处理后的第一数据的内存地址与所述第一内存中存储所述第一数据的内存地址相同。

3、如权利要求 1 或 2 所述的方法，其特征在于，所述目标数据承载于至少一个远程直接数据存取 RDMA 报文中，所述至少一个 RDMA 报文包括第一 RDMA 报文，所述第一 RDMA 报文为传输所述至少一个 RDMA 报文的过程中的最后一个 RDMA 报文，

所述第一设备的第一远程直接数据存取网络接口卡 RNIC 通过第二设备的第二 RNIC，从所述第二设备的第二内存中获取第一数据，包括：

所述第一 RNIC 接收所述第二 RNIC 发送的所述第一 RDMA 报文，所述第一 RDMA 报文中承载所述第一数据。

4、如权利要求 1 或 2 所述的方法，其特征在于，所述目标数据承载于多个 RDMA 报文中，所述多个 RDMA 报文包括第二 RDMA 报文和至少一个第三 RDMA 报文，所述第二 RDMA 报文为传输所述多个 RDMA 报文的过程中的最后一个 RDMA 报文，所述至少一个第三 RDMA 报文为传输所述多个 RDMA 报文的过程中传输所述最后一个 RDMA 报文之前的报文，

所述第一设备的第一远程直接数据存取网络接口卡 RNIC 通过第二设备的第二 RNIC，从所述第二设备的第二内存中获取第一数据，包括：

所述第一 RNIC 接收所述第二 RNIC 发送的第二 RDMA 报文和所述至少一个第三 RDMA 报文，所述第二 RDMA 报文中承载的数据和所述至少一个第三 RDMA 报文中承载的数据组成所述第一数据。

5、如权利要求 4 所述的方法，其特征在于，所述至少一个第三 RDMA 报文为一个 RDMA 报文，所述第三 RDMA 报文为传输所述多个 RDMA 报文的过程中传输所述最后一个 RDMA 报文的前一个 RDMA 报文。

6、一种数据校验的装置，其特征在于，包括：

获取单元，用于通过第二设备的第二远程直接数据存取网络接口卡 RNIC 从所述第二设备的第二内存中获取第一数据，所述第一数据为待传输的目标数据中的至少部分数据；

5 处理单元，用于将所述获取单元获取的所述第一数据缓存在第一设备的第一 RNIC 的缓存中；

所述处理单元，在所述第一 RNIC 将所述目标数据存储到所述第一设备的第一内存中之后，还用于从所述第一内存中读取所述目标数据中的第二数据，所述第二数据和所述第一数据为所述目标数据中位置相同的数据；

10 发送单元，若所述第一 RNIC 确定所述第一数据和所述第二数据相同，用于向所述第二 RNIC 发送确认字符 ACK。

7、如权利要求 6 所述的装置，其特征在于，

所述处理单元，还用于对所述第一数据进行数据处理得到处理后的第一数据，所述处理后的第一数据与所述第一数据不同；

15 所述处理单元，还用于将所述处理后的第一数据存储在第一内存中，所述第一内存中存储所述处理后的第一数据的内存地址与所述第一内存中存储所述第一数据的内存地址相同。

8、如权利要求 6 或 7 所述的装置，其特征在于，承载所述目标数据的报文包括至少一个远程直接数据存取 RDMA 报文，所述至少一个 RDMA 报文包括第一 RDMA 报文，所述第一 RDMA 报文为传输所述至少一个 RDMA 报文的过程中的最后一个 RDMA 报文，

20 所述获取单元，具体用于接收所述第二 RNIC 发送的所述第一 RDMA 报文，所述第一 RDMA 报文中承载所述第一数据。

9、如权利要求 6 或 7 所述的装置，其特征在于，承载所述目标数据的报文包括多个 RDMA 报文，所述多个 RDMA 报文包括第二 RDMA 报文和至少一个第三 RDMA 报文，所述第二 RDMA 报文为传输所述多个 RDMA 报文的过程中的最后一个 RDMA 报文，所述至少一个第三 RDMA 报文为传输所述多个 RDMA 报文的过程中传输所述最后一个 RDMA 报文之前的报文，

所述获取单元，具体用于接收所述第二 RNIC 发送的所述第二 RDMA 报文和所述至少一个第三 RDMA 报文，所述第二 RDMA 报文中承载的数据和所述至少一个第三 RDMA 报文中承载的数据组成所述第一数据。

30 10、如权利要求 9 所述的装置，其特征在于，所述至少一个第三 RDMA 报文为一个 RDMA 报文，所述第三 RDMA 报文为传输所述多个 RDMA 报文的过程中传输所述最后一个 RDMA 报文的前一个报文。

11、一种网卡，其特征在于，设置在第一设备中，所述网卡包括处理器、存储器以及收发器，所述存储器存储有程序指令，所述处理器运行所述程序指令以执行以下步骤：

35 控制所述收发器通过第二设备的第二网卡，从所述第二设备的第二内存中获取第一数据，所述第一数据为待传输的目标数据中的至少部分数据；

将所述第一数据缓存在所述网卡的缓存中；

40 将所述目标数据存储到所述第一设备的第一内存中之后，从所述第一内存中读取所述目标数据中的第二数据，所述第二数据和所述第一数据为所述目标数据中位置相同的数据；

若确定所述第一数据和所述第二数据相同，控制所述收发器向所述第二 RNIC 发送确认字符 ACK。

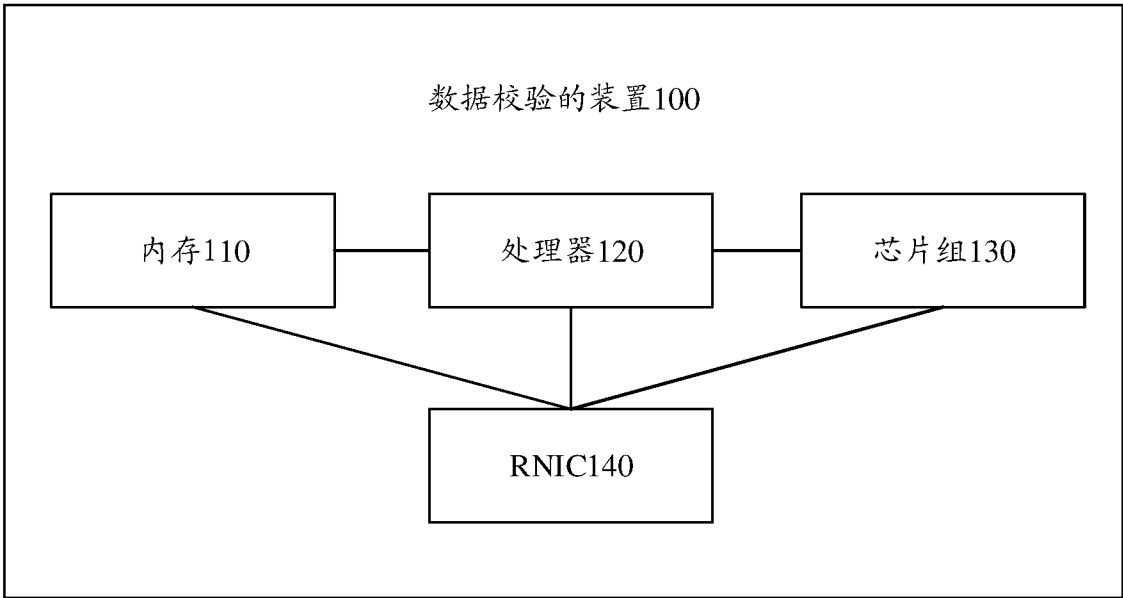


图 1

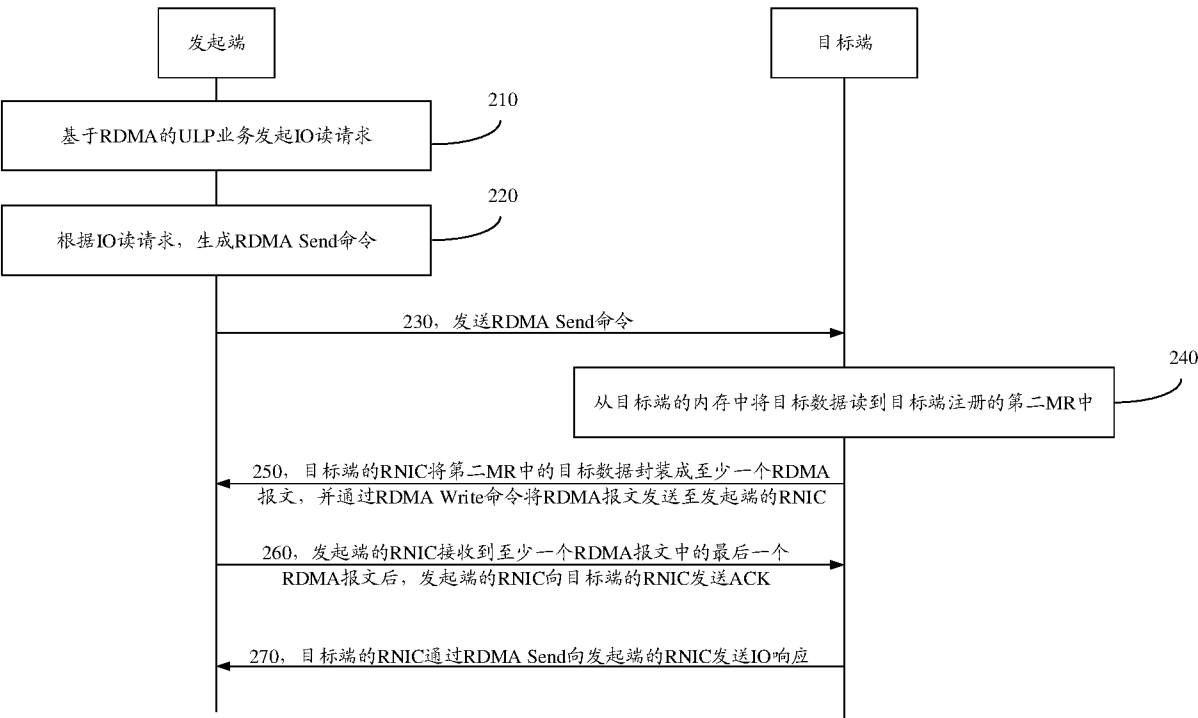


图 2

2/6

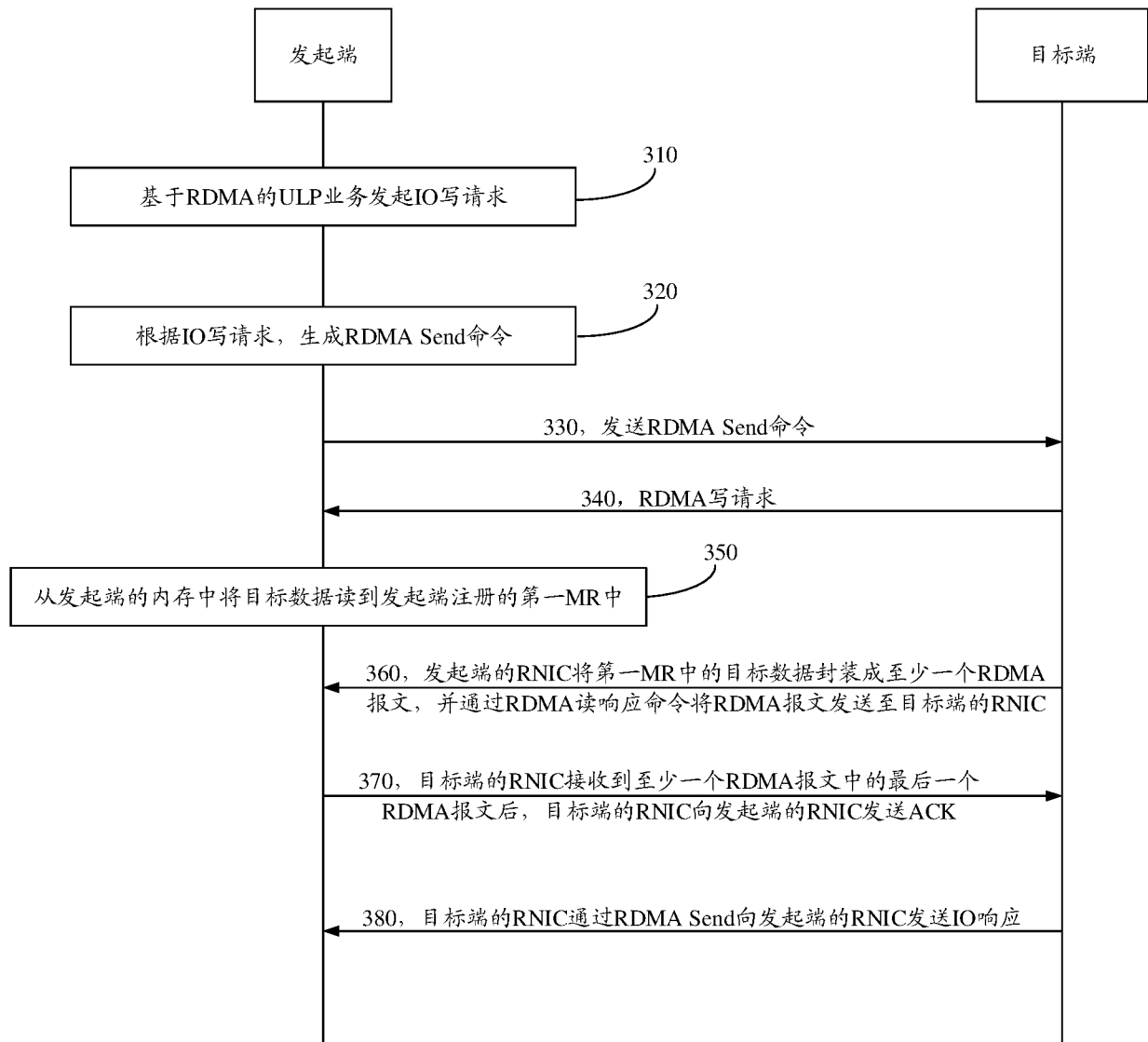


图 3

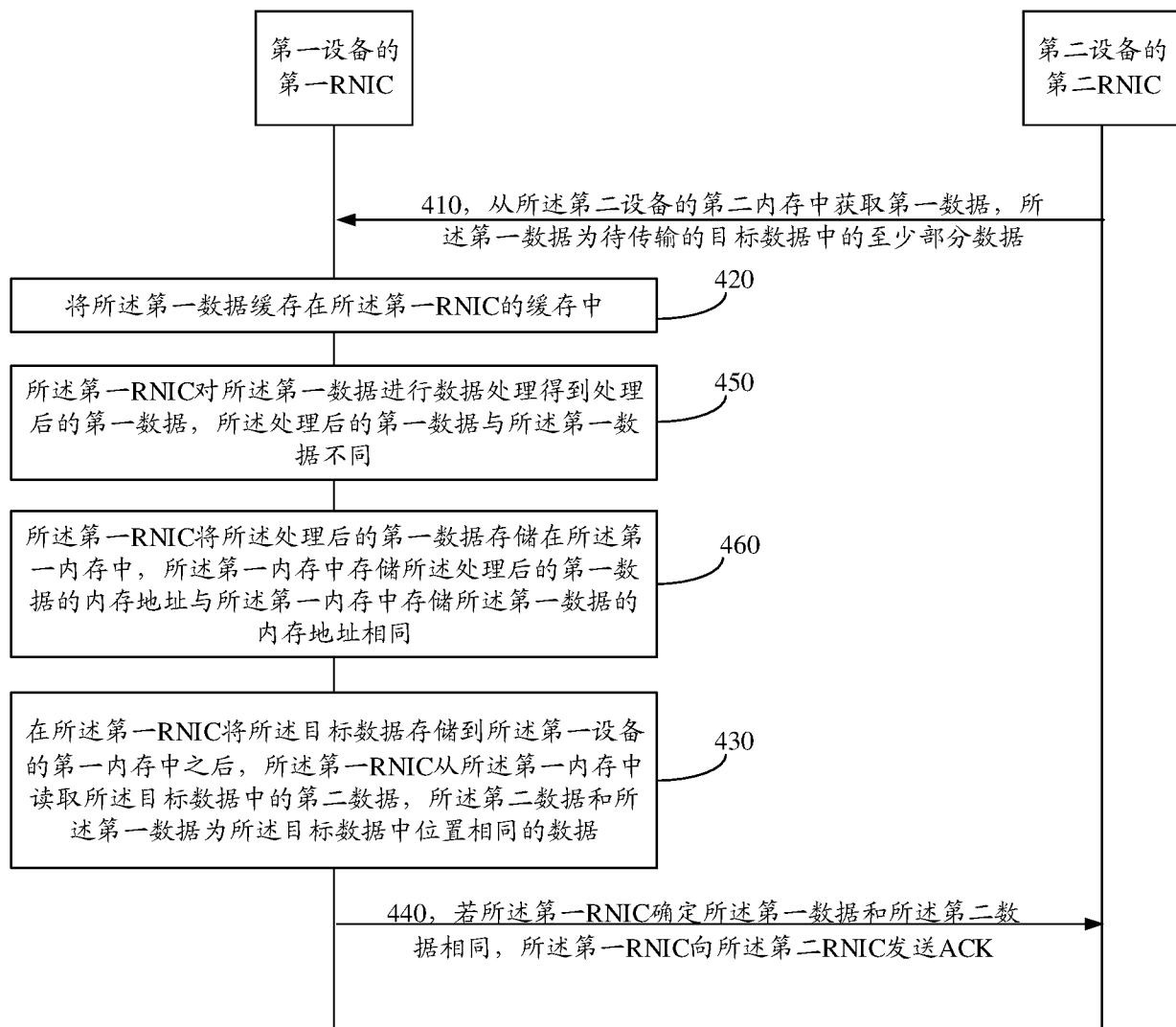


图 4

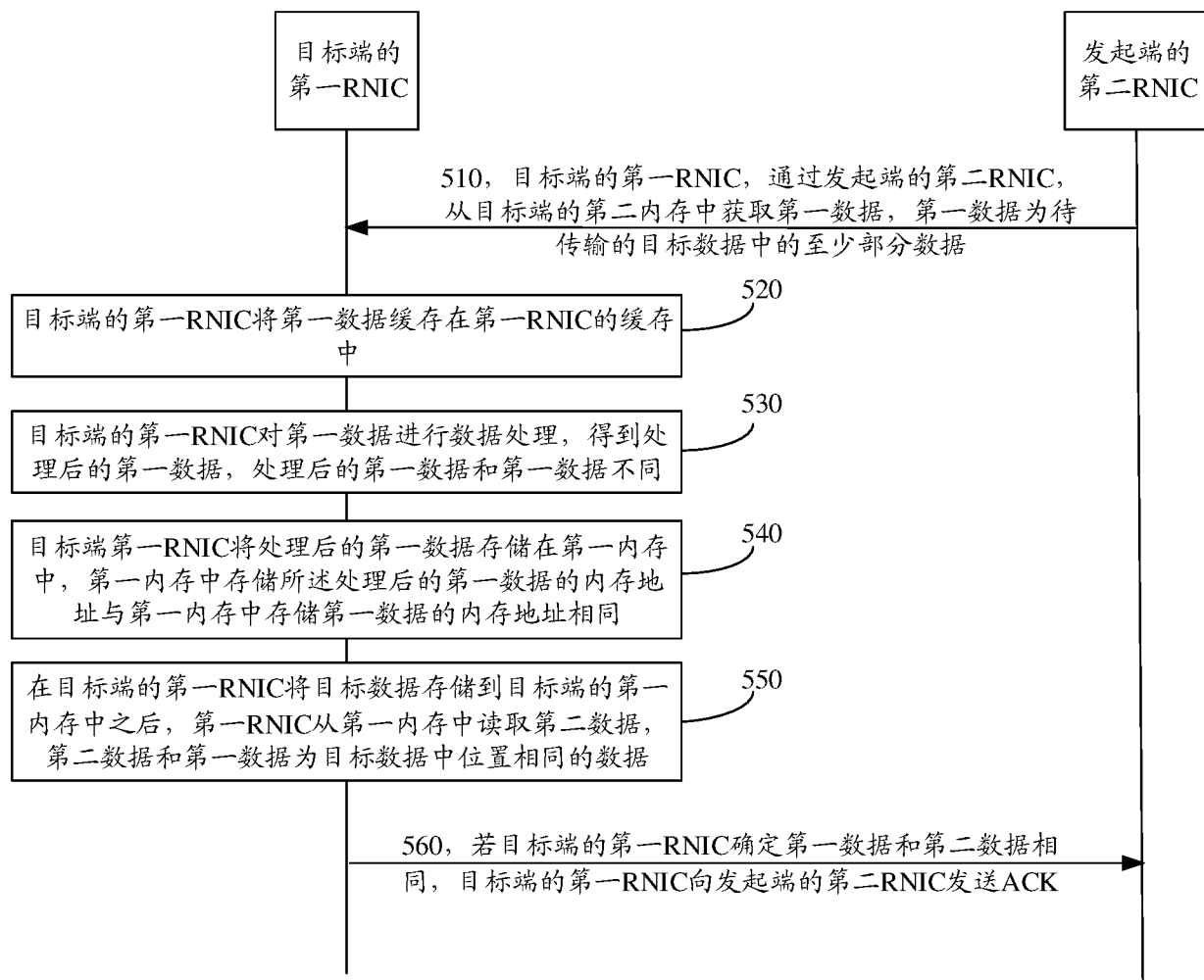


图 5

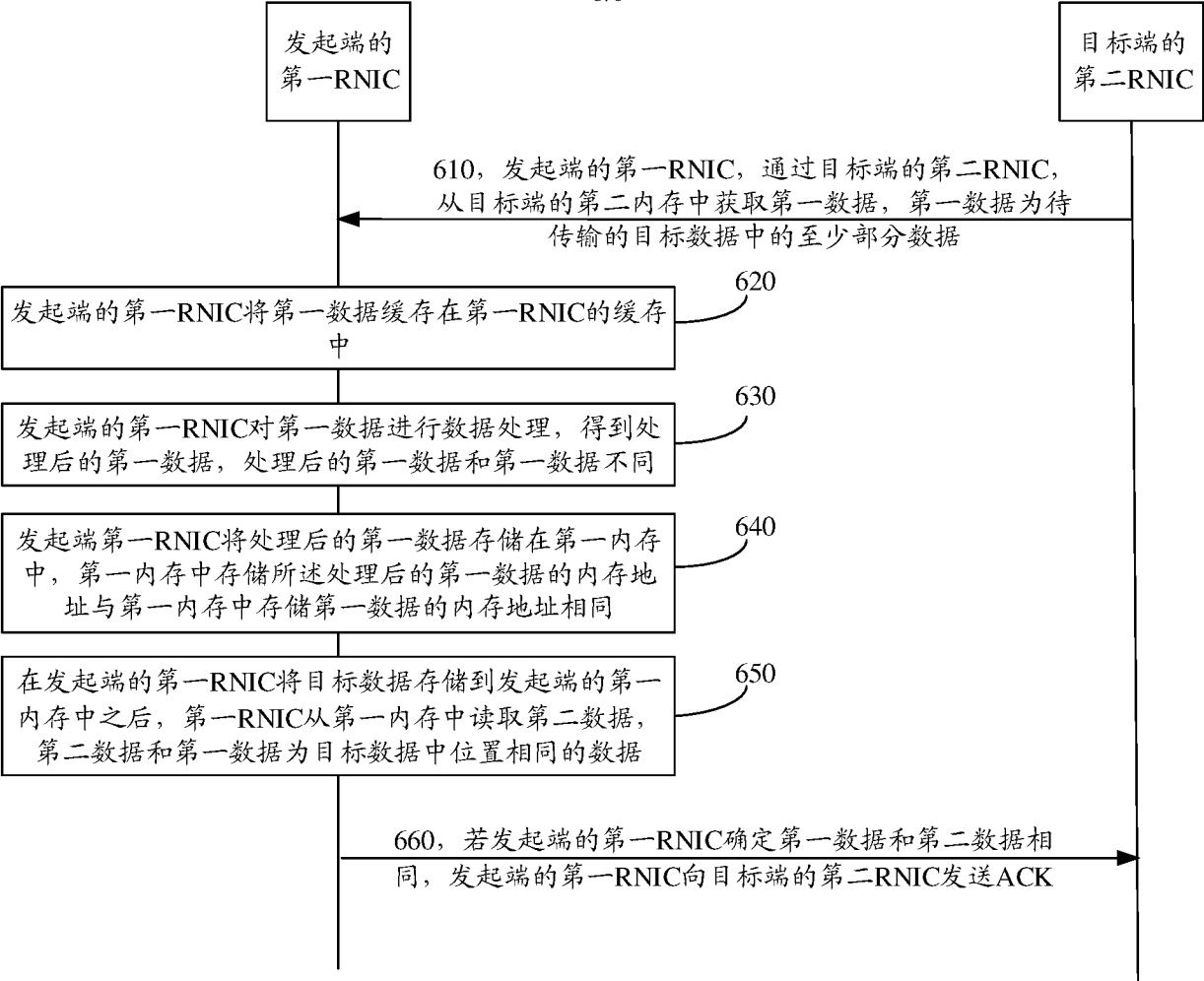


图 6

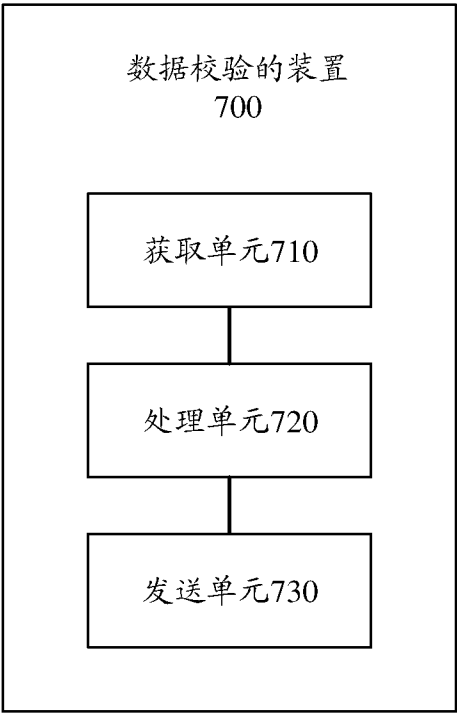


图 7

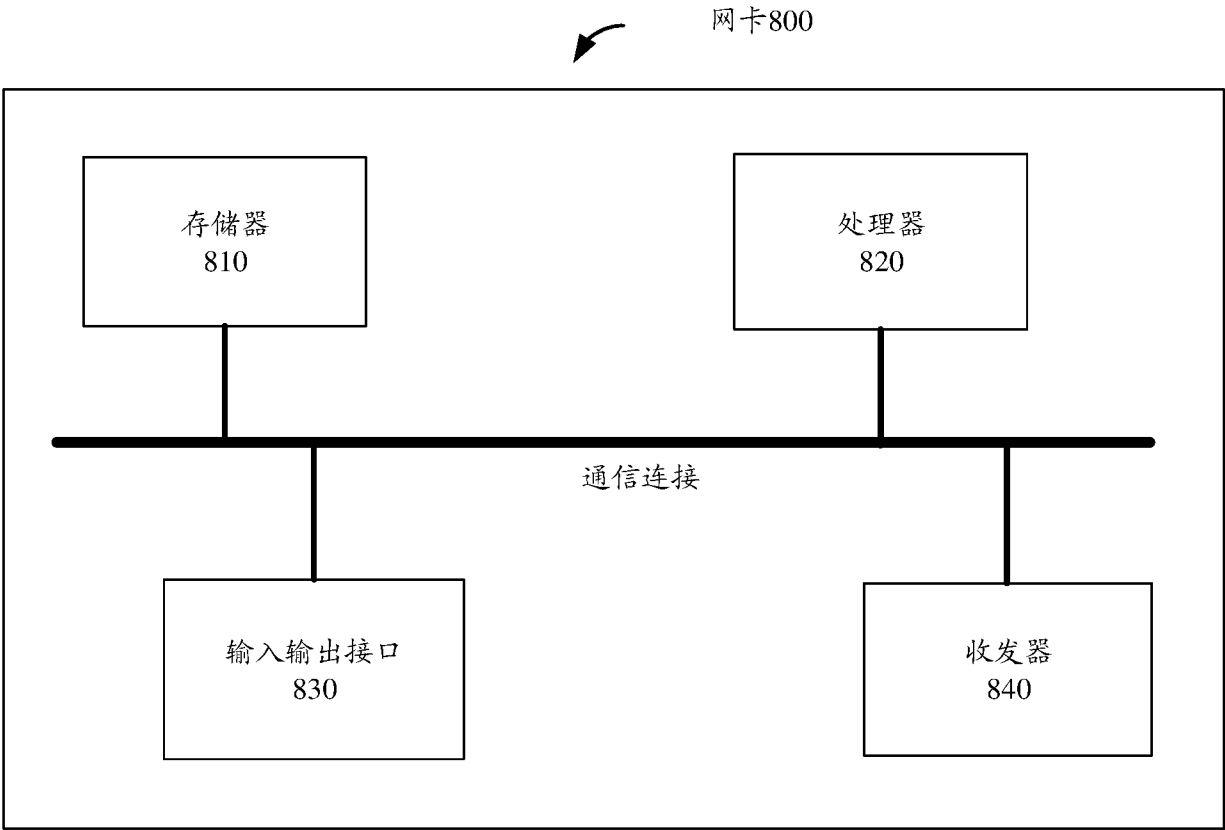


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/105957

A. CLASSIFICATION OF SUBJECT MATTER

G06F 13/14(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F; H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNKI, CNPAT, WPI, EPODOC, GOOGLE: 数据, data, packet, 部分, part, 目标, target, 有效, validity, 缓存, buffer, 传输, transmission, 远程, remote, 直接, direct, 内存, memory, 存取, access, RDMA, 网卡, RNIC, 读取, read, 校验, check, 验证, verify, 位置, location, 确认, ACK

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 107766261 A (HUAWEI TECHNOLOGIES CO., LTD.) 06 March 2018 (2018-03-06) claims 1-11	1-11
A	CN 103645994 A (HUAWEI TECHNOLOGIES CO., LTD.) 19 March 2014 (2014-03-19) claims 1-10	1-11
A	CN 101556573 A (HUAWEI SYMANTEC TECHNOLOGIES CO., LTD.) 14 October 2009 (2009-10-14) entire document	1-11
A	US 8365057 B2 (MELLANOX TECHNOLOGIES LTD.) 29 January 2013 (2013-01-29) entire document	1-11
A	US 7813369 B2 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 12 October 2010 (2010-10-12) entire document	1-11



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

16 November 2018

Date of mailing of the international search report

21 December 2018

Name and mailing address of the ISA/CN

State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/105957

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	107766261	A	06 March 2018	None			
CN	103645994	A	19 March 2014	None			
CN	101556573	A	14 October 2009	None			
US	8365057	B2	29 January 2013	US	2011029847	A1	03 February 2011
US	7813369	B2	12 October 2010	US	2006045108	A1	02 March 2006

国际检索报告

国际申请号

PCT/CN2018/105957

A. 主题的分类 G06F 13/14 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类																				
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06F; H04L 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNKI, CNPAT, WPI, EPODOC, GOOGLE: 数据, data, packet, 部分, part, 目标, target, 有效, validity, 缓存, buffer, 传输, transmission, 远程, remote, 直接, direct, 内存, memory, 存取, access, RDMA, 网卡, RNIC, 读取, read, 校验, check, 验证, verify, 位置, location, 确认, ACK																				
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 107766261 A (华为技术有限公司) 2018年 3月 6日 (2018 - 03 - 06) 权利要求1-11</td> <td>1-11</td> </tr> <tr> <td>A</td> <td>CN 103645994 A (华为技术有限公司) 2014年 3月 19日 (2014 - 03 - 19) 权利要求1-10</td> <td>1-11</td> </tr> <tr> <td>A</td> <td>CN 101556573 A (成都市华为赛门铁克科技有限公司) 2009年 10月 14日 (2009 - 10 - 14) 全文</td> <td>1-11</td> </tr> <tr> <td>A</td> <td>US 8365057 B2 (MELLANOX TECHNOLOGIES LTD.) 2013年 1月 29日 (2013 - 01 - 29) 全文</td> <td>1-11</td> </tr> <tr> <td>A</td> <td>US 7813369 B2 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2010年 10月 12日 (2010 - 10 - 12) 全文</td> <td>1-11</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 107766261 A (华为技术有限公司) 2018年 3月 6日 (2018 - 03 - 06) 权利要求1-11	1-11	A	CN 103645994 A (华为技术有限公司) 2014年 3月 19日 (2014 - 03 - 19) 权利要求1-10	1-11	A	CN 101556573 A (成都市华为赛门铁克科技有限公司) 2009年 10月 14日 (2009 - 10 - 14) 全文	1-11	A	US 8365057 B2 (MELLANOX TECHNOLOGIES LTD.) 2013年 1月 29日 (2013 - 01 - 29) 全文	1-11	A	US 7813369 B2 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2010年 10月 12日 (2010 - 10 - 12) 全文	1-11
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 107766261 A (华为技术有限公司) 2018年 3月 6日 (2018 - 03 - 06) 权利要求1-11	1-11																		
A	CN 103645994 A (华为技术有限公司) 2014年 3月 19日 (2014 - 03 - 19) 权利要求1-10	1-11																		
A	CN 101556573 A (成都市华为赛门铁克科技有限公司) 2009年 10月 14日 (2009 - 10 - 14) 全文	1-11																		
A	US 8365057 B2 (MELLANOX TECHNOLOGIES LTD.) 2013年 1月 29日 (2013 - 01 - 29) 全文	1-11																		
A	US 7813369 B2 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2010年 10月 12日 (2010 - 10 - 12) 全文	1-11																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																				
国际检索实际完成的日期 2018年 11月 16日		国际检索报告邮寄日期 2018年 12月 21日																		
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10) 62019451		授权官员 刘承恩 电话号码 86-(10)-53961634																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/105957

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	107766261	A	2018年 3月 6日	无			
CN	103645994	A	2014年 3月 19日	无			
CN	101556573	A	2009年 10月 14日	无			
US	8365057	B2	2013年 1月 29日	US	2011029847	A1	2011年 2月 3日
US	7813369	B2	2010年 10月 12日	US	2006045108	A1	2006年 3月 2日