(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

(43) 国际公布日
2019 年 2 月 7 日 (07.02.2019)

WIPO | PCT

(10) 国际公布号
WO 2019/024763 A1

(51) 国际专利分类号:
H04L 12/861 (2013.01)

(21) 国际申请号: PCT/CN2018/097254

(22) 国际申请日: 2018 年 7 月 26 日 (26.07.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710638401.3 2017年7月31日 (31.07.2017) CN

(71) 申请人: 新 华 三 技 术 有 限 公 司 **(NEW H3C TECHNOLOGIES CO., LTD.)** [CN/CN]; 中国浙江省杭州市滨江区长河路 466 号, Zhejiang 310052 (CN)。

(72) 发明人: 田 浩 **(TIAN, Hao)**; 中国北京市海淀区上地信息产业基地创业路 2 号东方电子大厦211室, Beijing 100085 (CN)。

(74) 代理人: 北 京 博 思 佳 知 识 产 权 代 理 有 限 公 司 **(BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION)**; 中国北京市海淀区上地三街 9 号嘉华大厦 B 座 409, Beijing 100085 (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。

(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,

(54) **Title:** MESSAGE PROCESSING

(54) 发明名称: 报文处理

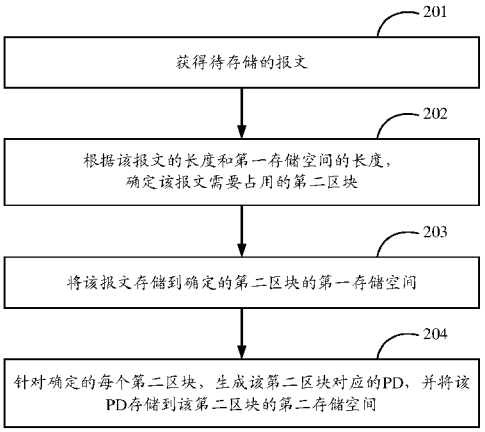


图 2A

201 Obtain a message to be stored
202 Determine, according to the length of the message and the length of a first storage space, a second block requiring occupation by the message
203 Store the message to the first storage space of the determined second block
204 Generate a PD corresponding to the second block for each determined second block, and store the PD to a second storage space of the second block

(57) **Abstract:** A memory of a network device is divided into a plurality of first blocks. Each of the first blocks is divided into a plurality of second blocks and each of the second blocks comprises a first storage space and a second storage space. When a message is stored, a second block requiring occupation by the message is determined according to the length of the message and the length of the first storage space, and the message is stored in the first storage space of the determined second block. A PD corresponding to the second block is generated for each determined second block, and the PD is stored in the second storage space of the second block. When the message is loaded, a second block to be loaded is determined according to a start address of the message to be loaded, and message segments are loaded from the first storage space of the determined second block. A message is composed from the loaded message segments, and the composed message is transmitted.

WO 2019/024763 A1

[见续页]

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))。

(57) 摘要: 网络设备的内存被划分成多个第一区块，每个第一区块被划分成多个第二区块，每个第二区块均包括第一存储空间和第二存储空间。在存储报文时，根据报文的长度和第一存储空间的长度，确定报文需要占用的第二区块，并将所述报文存储到所确定的第二区块的第一存储空间。针对所确定的每个第二区块，生成该第二区块对应的PD，并将所述PD存储到该第二区块的第二存储空间。在读取报文时，根据待读取报文的开始地址确定要读取的第二区块，并从所确定的第二区块的第一存储空间读取报文片段，将所读取的报文片段组成报文，并发送组成的报文。

报文处理

相关申请的交叉引用

[01]本专利申请要求于2017年7月31日提交的、申请号为201710638401.3、发明名称为“一种报文处理方法、装置和网络设备”的中国专利申请的优先权，该申请的全文以引用的方式并入本文中。

背景技术

[02]网络设备在接收到报文时，可根据报文所属的队列，将报文缓存到外部的存储单元中。网络设备在发送报文时，也可根据报文所属的队列，从外部的存储单元中读出报文并发送。在报文的存取过程中存在两个关键元素，分别为PKT（PACKET，报文实体）和PD（Packet Description，报文描述符）。

[03]网络设备可将PKT存储到DDR SDRAM（Dual Data Rate Synchronous Dynamic Random Access Memory，双倍速率同步动态随机存取存储器，后续简称DDR），并将PD存储到QDR SRAM（Quad Data Rate Static Random Access Memory，四倍数据速率静态随机存取存储器，后续简称QDR）。由于使用QDR存储PD，提高了单板设计成本，增加了布线压力。此外，在发送报文时，网络设备先从QDR读取PD，然后利用PD从DDR读取PKT。由于分两次分别读取PD和PKT，读取效率较低。

附图说明

[04]图1A-图1C是本申请一种实施方式中的应用场景示意图。

[05]图2A是本申请一种实施方式中的报文存储的示意性流程图。

[06]图2B是本申请一种实施方式中的确定报文需要占用的第二区块的示意性流程图。

[07]图2C是本申请一种实施方式中的报文读取的示意性流程图。

[08]图3A-图3E是本申请另一种实施方式中的应用场景示意图。

[09]图4是本申请一种实施方式中的报文处理装置的结构图。

[10]图5A和图5B是本申请一种实施方式中的网络设备的硬件结构图。

具体实施方式

[11]在本申请实施例使用的术语仅仅是出于描述特定实施例的目的，而非限制本申请。本申请和权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其它含义。还应当理解，本文中使用的术语“和/或”是指包含一个或多个相关联的列出项目的任何或所有可能组合。

[12]应当理解，尽管在本申请实施例可能采用术语第一、第二、第三等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本申请范围的情况下，第一信息也可以被称为第二信息，类似地，第二信息也可以被称为第一信息。取决于语境，此外，所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[13]本申请实施例中提出一种报文处理方法，可以应用于网络设备。网络设备在处理报文时，会涉及报文的存储和读取过程。在报文的存取过程中，会存在两个关键元素，分别为待存储的报文本身和 PD。待存储的报文本身也可以称为 PKT，以下也可简称为报文；而 PD 用于记录报文的关键信息，如报文的长度、队列信息、存储地址等。PD 的长度远小于报文的长度，PD 的信息量比较小。

[14]本申请实施例中，将报文和 PD 存储到同一个内存、如 DDR 内存，以充分利用 DDR 内存资源，从而有效降低单板设计成本、减少布线压力。而且，由于将报文和 PD 存储到同一个内存，在需要发送报文时，可以从同一个内存中读取到报文和 PD，读取效率相对很高。

[15]本申请实施例中，内存（如 DDR 内存）被划分成多个第一区块，每个第一区块又被划分成多个第二区块，每个第二区块可以为一个 burst。每个第二区块又被划分成为大小不等的两段存储空间，即第一存储空间和第二存储空间，其中，第一存储空间可以大于第二存储空间。基于此，可以将报文存储到第一存储空间，将 PD 存储到第二存储空间，从而将报文和 PD 存储到同一个内存。

[16]在一个例子中，参见图 1A 所示，DDR 内存可以被划分成了多个第一区块，每个第一区块为 2K 字节的大块内存。此外，每个第一区块可以被划分成 32 个第二区块，每个第二区块为 64 字节的小块内存，即每个 2K 字节的第一区块被划分成 32 个 64 字节的第二区块，每一个第二区块可以为一个 burst。

[17]其中，上述 2K 字节、32 个、64 字节等，只是本申请的一个示例，对此数值不做限

制。例如，假设第一区块为 2K 字节，第二区块为 128 字节时，则第一区块可以被划分成 16 个第二区块。又例如，假设第一区块为 4K 字节，第二区块为 64 字节时，则第一区块可以被划分成 64 个第二区块。以此类推。为了方便描述，后续以 2K 字节的第一区块被划分成 32 个 64 字节的第二区块为例。

[18]其中，在将第一区块划分成 32 个第二区块之后，还可以将每个第二区块划分为大小不等的两段存储空间。较大的第一存储空间可以为例如 56 字节，第一存储空间用于存储报文。较小的第二存储空间可以为例如 8 字节，第二存储空间用于存储 PD。参见图 1B 所示，第二区块可以同时存储报文和 PD。例如，假设需要存储 4 个报文（如报文 0、报文 1、报文 2、报文 3），且 4 个报文的大小均是 96 字节，则报文 0 存储在第二区块 0 和第二区块 1，报文 0 可以占用第二区块 0 的第一存储空间 56 字节，占用第二区块 1 的第一存储空间内 40 字节，且在第二区块 0 的前 8 字节的第二存储空间中存储有 PD，在第二区块 1 的前 8 字节的第二存储空间中也存储有 PD。以此类推。

[19]参见表 1，为 PD 所包括的信息域的一个示例。

信息域	信息域所表示的信息
LEN	报文的长度，存储在该 PD 对应的第二区块中
SOP	该 PD 对应的第二区块是否为所存储报文占用的第一个第二区块
EOP	该 PD 对应的第二区块是否为所存储报文占用的最后一个第二区块
BV	该 PD 对应的第二区块是否有效
VAL	该 PD 对应的第二区块是否为该第二区块所在的第一区块内的最后一个第二区块
NBA	指向该 PD 对应的第二区块所存储的报文占用的下一个第一区块
NBN	该 PD 对应的第二区块所存储的报文在下一个第一区块中占用的第二区块数量

表 1

[20]参见图 1C 所示，为 PD 的一个示例。为了简化视图，图 1C 中仅示出了 PD 中存在有效值的部分信息域。不同第二区块中存储的 PD 可能相同，也可能不同。例如，当一个报文占用一个或多个第二区块时，对于占用的所有第二区块中的第一个第二区块，则需要对该第二区块中的 PD 内的信息域 LEN 和 SOP 赋值。例如，当某个第二区块中存

储的PD内的信息域SOP的取值为0时，表示该第二区块是存储报文的第一个第二区块。当信息域SOP（下文中也可称为第一标记）的取值为1时，表示该第二区块不是存储报文的第一个第二区块。信息域LEN的取值为该PD对应的第二区块所存储的报文的长度。

[21]对于报文需要占用的所有第二区块中的最后一个第二区块，则需要对该第二区块中的PD内的信息域EOP（下文中也可称为第二标记）赋值。例如，当某个第二区块中存储的PD内的信息域EOP的取值为0时，表示该第二区块是存储报文的最后一个第二区块。当信息域EOP的取值为1时，表示该第二区块不是存储报文的最后一个第二区块。

[22]对于报文需要占用的任一第二区块，则需要将该第二区块中的PD内的信息域BV赋值为有效。例如，当某个第二区块中存储的PD内的信息域BV的取值为0时，表示该第二区块有效。然而，当信息域BV的取值为1时，表示该第二区块无效。

[23]此外，当一个报文需要占用的多个第二区块属于至少两个第一区块时，若第二区块是该报文需要占用的当前第一区块中的最后一个第二区块、但不是该报文需要占用的最后一个第二区块，则需要对该第二区块中的PD内的信息域VAL、NBA和NBN赋值。其中，信息域VAL可以表示信息域NBA是否有效。例如，当某个第二区块中存储的PD内的信息域VAL的取值为0时，则表示该PD内的信息域NBA有效；然而，当该PD内的信息域VAL的取值为1时，则表示该PD内的信息域NBA无效。信息域NBA（Next Block Address，下一块地址）的取值为该报文需要占用的下一个第一区块的指针地址。信息域NBN（Next Block Number，下一块数量）的取值为该报文在下一个第一区块中占用的第二区块数量。

[24]例如，在图1C中，报文1占用第二区块2、第二区块3和第二区块4。报文1需要占用的所有第二区块中的第一个第二区块为第二区块2，该第二区块2对应的PD内的信息域LEN的取值为例如120，以表示报文1的总长度是120字节。该第二区块2中的PD内的信息域SOP的取值为0，以表示该第二区块2为报文1占用的第一个第二区块。针对报文1需要占用的任意一个第二区块（如第二区块2、第二区块3和第二区块4，图1C中为简化视图仅以第二区块3为例示出了BV），该第二区块3对应的PD内的信息域BV的取值为0，以表示该第二区块3有效。报文1需要占用的所有第二区块中的最后一个第二区块为第二区块4，该第二区块4中的PD内的信息域EOP的取值为0，以表示该第二区块4为报文1占用的最后一个第二区块。

[25]在这个示例中，第二区块2中的PD可以记录：LEN=120；SOP=0；BV=0；第二区块4中的PD可以记录：EOP=0；BV=0。第二区块3中的PD可以记录：BV=0。

[26]又例如，报文 2 占用第一区块 0 中的第二区块 5 至第二区块 31 以及第一区块 1 中的第二区块 0、第二区块 1。报文 2 需要占用的所有第二区块中的第一个第二区块为第一区块 0 中的第二区块 5，该第一区块 0 中的第二区块 5 对应的 PD 内的信息域 LEN 的取值为例如 1560，以表示报文 2 的总长度是 1560 字节。该第一区块 0 中的第二区块 5 中的 PD 内的信息域 SOP 的取值为 0，以表示该第二区块 5 为报文 2 占用的第一个第二区块。针对报文 2 需要占用的任意一个第二区块（如第一区块 0 中的第二区块 5 至第二区块 31、以及第一区块 1 中的第二区块 0 和第二区块 1），这些第二区块对应的 PD 内的信息域 BV 的取值均为 0，以表示这些第二区块有效。报文 2 占用的第一区块 0 中的第二区块 31 为第一区块 0 中的最后一个第二区块，该第一区块 0 中的第二区块 31 对应的 PD 中的信息域 NBA 记录了报文 2 占用的下一个第一区块、即第一区块 1 的指针，如第一区块 1 的第二区块 0 的地址。第一区块 0 中的第二区块 31 对应的 PD 中的信息域 NBN 的取值为 2，以记录报文 2 在第一区块 1 中占用的第二区块数量。第一区块 0 中的第二区块 31 对应的 PD 内的信息域 VAL 的取值为 0，以表示信息域 NBA 有效。报文 2 需要占用的所有第二区块中的最后一个第二区块为第一区块 1 中的第二区块 1，该第一区块 1 中的第二区块 1 对应的 PD 内的信息域 EOP 取值为 0，以表示该第一区块 1 中的第二区块 1 为报文 2 占用的最后一个第二区块。

[27]在这个示例中，第一区块 0 中的第二区块 5 的 PD 可以记录：LEN=1560；SOP=0；BV=0；第一区块 0 中的第二区块 31 的 PD 可以记录：BV=0；NBA=0x8049324（其表示占用的下一个第一区块的指针地址）；NBN=2；VAL=0；对于其它第二区块的 PD，在此不再赘述。

[28]在上述应用场景下，参见图 2A 所示，为报文存储的示意性流程图。

[29]步骤 201，获得待存储的报文。

[30]步骤 202，根据该报文的长度和第一存储空间的长度，确定该报文需要占用的一个或多个第二区块。

[31]步骤 203，将该报文存储到确定的第二区块的第一存储空间。

[32]步骤 204，针对确定的每个第二区块，生成该第二区块对应的 PD，并将该 PD 存储到该第二区块的第二存储空间。

[33]在一个例子中，如图 2B 所示，在步骤 202“确定该报文需要占用的一个或多个第二区块”可以包括但不限于如下方式：

[34]步骤2021，确定该报文对应的队列。

[35]步骤2022，判断该队列是否对应有尾指针。如果有，则进入步骤2023，否则进入步骤2024。

[36]步骤2023，若该队列对应有尾指针，则从该尾指针指向的第二区块开始，确定该报文需要占用的第二区块；

[37]步骤2024，若该队列没有对应尾指针，则可以选取空闲的第一区块，并从选取的第一区块中的首个第二区块开始，确定该报文需要占用的第二区块。

[38]其中，可以先根据该报文的长度和第一存储空间的长度，确定该报文需要占用的第二区块的数量N。然后，从该尾指针指向的第二区块开始或者从选取的第一区块中的首个第二区块开始，确定该报文需要占用的N个第二区块。每个第二区块用于存储该报文的一个报文片段，例如，若报文需要占用3个第二区块，则可以将报文划分为3个报文片段，每个报文片段占用一个第二区块。

[39]其中，网络设备可以采用队列存储报文，每个队列具有唯一的队列标识。当然，采用队列存储报文的方式只是一个示例，网络设备还可以采用其它方式存储报文。通过队列存储报文，可以有效的将不同类型的报文存储到不同的队列。例如，网络设备包括三个队列（如FIFO（First-In First-Out，先进先出）队列），这三个队列的队列标识分别为队列A、队列B和队列C。其中，队列A用于存储具有特征A（如出接口为接口A）的报文，队列B用于存储具有特征B（如出接口为接口B）的报文，队列C用于存储具有特征C（如出接口为接口C）的报文。

[40]网络设备在获得待存储的报文之后，可以先确定该报文对应的队列。若该报文是该队列对应的首个报文，则该队列没有对应尾指针和头指针，若该报文不是该队列对应的首个报文，则该队列对应有尾指针和头指针。

[41]其中，头指针表示指向队列对应的第一个报文头部的指针，如队列中第一个报文的开始地址或者开始位置。尾指针表示指向队列对应的最后一个报文尾部的指针，如队列中最后一个报文的结束地址或者结束位置。

[42]例如，在队列为空的情况下，若获得待存储的报文1，则在队列中存储报文1，并更新该队列的头指针为指向报文1的开始地址，更新该队列的尾指针为指向报文1的结束地址。然后，若获得待存储到该队列的报文2，则在该队列中存储报文2，并更新该队列的尾指针为指向报文2的结束地址。以此类推，每接收到该队列的报文，就可以更新

该队列的尾指针。

[43]若需要读取队列中的报文，则从队列的头指针开始读取。例如，在报文 1 读取完成后，则从该队列中删除报文 1，并更新该队列的头指针为指向报文 2 的开始地址。以此类推，每从队列中读取并删除一个报文，就可以更新该队列的头指针。综上所述，存储报文的过程，是从队列的尾指针开始依次存储报文，读取报文的过程，是从队列的头指针开始依次读取报文。

[44]其中，步骤 2023 可以包括：从该尾指针指向的第二区块的下一个第二区块开始，选取 N 个第二区块，所述 N 是根据报文的长度以及第一存储空间的长度确定的。例如，若尾指针指向第一区块 1 中的第二区块 10，且根据报文的长度 100 字节、第一存储空间的长度 56 字节确定该报文需要占用两个第二区块，则该报文要占用的第二区块为第一区块 1 中的第二区块 11 和第二区块 12。又例如，若尾指针指向第一区块 1 中的第二区块 30，且根据报文的长度 100 字节、第一存储空间的长度 56 字节确定该报文需要占用两个第二区块，则由于第一区块 1 已经不存在两个空闲的第二区块，从而重新申请空闲的第一区块、如第一区块 3。基于此，该报文要占用的第二区块可以为第一区块 1 中的第二区块 31 和第一区块 3 中的第二区块 0。

[45]此外，在步骤 2023 中，在确定报文需要占用的第二区块之后，还可以更新队列的尾指针为指向该报文占用的最后一个第二区块的结束地址，如指向第二区块 12 的结束地址。

[46]其中，步骤 2024 可以包括：从选取的空闲的第一区块中的首个第二区块开始，选取 N 个第二区块，所述 N 是根据报文的长度以及第一存储空间的长度确定的。例如，若选取空闲的第一区块 2，且根据报文的长度 100 字节、第一存储空间的长度 56 字节确定该报文需要占用两个第二区块，则该报文要占用的第二区块可以为第一区块 2 中的第二区块 0 和第二区块 1。

[47]此外，在步骤 2024 中，在确定该报文需要占用的第二区块之后，还可以更新队列的头指针为指向该报文要占用的第一个第二区块的开始地址，以及更新队列的尾指针为指向该报文要占用的最后一个第二区块的结束地址。

[48]网络设备可以维护空闲地址池，该空闲地址池用于存储所有空闲的第一区块。在申请空闲区块时，可以以第一区块为单位从空闲地址池中选取该空闲区块，即每次需要申请一个或者多个第一区块，而不是直接申请一个或者多个第二区块。在回收内存时，也

是以第一区块为单位回收，即将一个或者多个第一区块回收到空闲地址池，而不是将一个或者多个第二区块回收到空闲地址池。

[49]在一个例子中，在步骤 203“将报文存储到确定的第二区块的第一存储空间”时，若报文需要占用 N 个第二区块，则可以将报文划分为 N 个报文片段，并可以将这 N 个报文片段分别存储到 N 个第二区块的第一存储空间。例如，以报文的长度是 100 字节、报文需要占用第二区块 11 和第二区块 12 为例，则可以将该报文的前 56 个字节划分到报文片段 1，并将该报文片段 1 存储到该第二区块 11 的第一存储空间；将该报文的后 44 个字节划分到报文片段 2，并将该报文片段 2 存储到该第二区块 12 的第一存储空间。

[50]在一个例子中，当报文需要占用一个或多个第二区块时，在步骤 204“生成该第二区块对应的 PD”可以包括但不限于：若该第二区块是报文需要占用的所有第二区块中的第一个第二区块，则需要在该第二区块的 PD 中记录该报文的长度，以及将该第二区块的 PD 内的第一标记置为有效，以表示该第二区块为报文占用的第一个第二区块。若该第二区块是报文需要占用的所有第二区块中的最后一个第二区块，则需要将该第二区块的 PD 内的第二标记置为有效，以表示该第二区块为报文占用的最后一个第二区块。若第二区块是报文需要占用的任一第二区块，则需要将该第二区块的 PD 内的指示当前区块是否被占用的标记置为有效。在本公开示例中，PD 内的任意标记均可通过赋值的形式来被设置。例如，可以对信息域 SOP、信息域 EOP、信息域 BV 等等分别进行赋值。

[51]此外，当报文需要占用的所有第二区块属于至少两个第一区块时，若该第二区块是报文在某个第一区块中占用的最后一个第二区块，则需要将该第二区块的 PD 内的信息域 NBA 赋值为指向报文需要占用的下一个第一区块的指针，信息域 NBN 赋值为报文在下一个第一区块中需要占用的第二区块数量、信息域 VAL 赋值为有效。

[52]上述 PD 的生成过程可以参见图 1C 所示，在此不再重新赘述。

[53]在一个例子中，在步骤 204“将该 PD 存储到该第二区块的第二存储空间”时，假设报文需要占用第二区块 11 和第二区块 12，则确定要将报文长度存储到第二区块 11 的第二存储空间中对应 PD 的信息域 LEN 的位置处，对第二区块 11 的第二存储空间中对应 PD 的信息域 SOP 赋值以表示第二区块 11 为报文占用的第一个第二区块，并对第二区块 11 的第二存储空间中对应 PD 的信息域 BV 赋值以表示第二区块 11 被报文占用；确定要对第二区块 12 的第二存储空间中对应 PD 的信息域 EOP 赋值以表示第二区块 12 为报文占用的最后一个第二区块，并对并对第二区块 12 的第二存储空间中对应 PD 的信息域 BV 赋值以表示第二区块 12 被报文占用。

[54]基于上述技术方案，通过将 DDR 内存划分成多个第一区块，每个第一区块被划分成多个第二区块，每个第二区块包括第一存储空间和第二存储空间，可以将报文存储到第一存储空间，并将 PD 存储到第二存储空间。这样，可以将报文和 PD 存储到同一个内存，可充分利用 DDR 资源，并有效降低了单板设计成本，减少了布线压力。

[55]相应地，参见图 2C 所示，为报文读取的示意性流程图，包括:

[56]步骤 210，根据要读取的报文的开始地址确定要读取的一个或多个第二区块。

[57]步骤 211，从要读取的各第二区块的第一存储空间读取报文片段。

[58]步骤 212，从该第二区块的第二存储空间读取 PD。

[59]步骤 213，利用读取的 PD 将读取的报文片段组成报文。

[60]步骤 214，发送组成的报文。

[61]其中，图 2C 所示步骤 210 至步骤 214 是报文的读取过程，以及报文的发送过程。

[62]具体的，从每个要读取的第二区块的第二存储空间读取 PD，并将信息域 SOP 有效的 PD 对应的第二区块确定为起始第二区块，将信息域 EOP 有效的 PD 对应的第二区块确定为终止第二区块。若从当前的第一区块未读取到信息域 EOP 有效的 PD，则从所述第一区块中的最后一个第二区块的第二存储空间读取 PD，以根据该 PD 中的信息域 NBA 获取指向下一个第一区块的指针、根据该 PD 中的信息域 NBN 获取报文在下一个第一区块中占用的第二区块数量，并根据所述指针和所述第二区块数量，从下一个第一区块中的第二区块的第二存储空间读取 PD。以此类推，直至读取到信息域 EOP 有效的 PD 为止，并将信息域 EOP 有效的 PD 对应的第二区块确定为终止第二区块。从起始第二区块与终止第二区块之间的每个第二区块的第一存储空间读取报文片段，并将所读取的报文片段组成报文，并发送所组成的报文。

[63]在需要发送某个队列中的报文时，查询出该队列的头指针，这个头指针指向该队列中的第一个报文的头部、例如第一个报文的开始地址，且这个开始地址可以是第二区块的地址。然后，确定该第二区块所属的第一区块，并从该第一区块的该第二区块的第二存储空间读取 PD。

[64]参见图 1C 所示，由于第一区块 0 中第二区块 0 的 PD 内的信息域 SOP 有效，第一区块 0 中第二区块 1 的 PD 内的信息域 EOP 有效，因此将第一区块 0 中第二区块 0 确定为起始第二区块，并将第一区块 0 中第二区块 1 确定为终止第二区块。相应地，从第一

区块 0 中第二区块 0、第二区块 1 的第一存储空间中读取报文片段，并将读取的报文片段组成报文 0，并发送该报文 0。

[65]由于第一区块 0 中第二区块 2 的 PD 内的信息域 SOP 有效，第一区块 0 中第二区块 4 的 PD 内的信息域 EOP 有效，因此将第一区块 0 中第二区块 2 确定为起始第二区块，将第一区块 0 中第二区块 4 确定为终止第二区块。相应地，从第一区块 0 中的第二区块 2、第二区块 3、第二区块 4 的第一存储空间中读取报文片段，并将读取的报文片段组成报文 1，并发送该报文 1。

[66]又例如，由于第一区块 0 中第二区块 5 的 PD 内的信息域 SOP 有效，因此将第一区块 0 中第二区块 5 确定为起始第二区块。由于在第一区块 0 中，第二区块 5 后的各第二区块 6-31 的 PD 内的信息域 EOP 均不是有效，因此从第二区块 31 的 PD 中的信息域 NBA 获取指向下一个第一区块的指针、从第二区块 31 的 PD 中的信息域 NBN 获取报文在下一个第一区块占用的第二区块数量 M，并根据获取到的指针确定下一个第一区块的地址（图中示例为第一区块 1），从第一区块 1 中前 M 个第二区块（图中示例为第二区块 0 和第二区块 1）的第二存储空间读取 PD。由于第一区块 1 中的第二区块 1 的 PD 内的信息域 EOP 有效，因此将第一区块 1 中的第二区块 1 确定为终止第二区块。然后，从第一区块 0 中的第二区块 5-第二区块 31、第一区块 1 中的第二区块 0 和第二区块 1 的第一存储空间读取报文片段，并将读取的报文片段组成报文 2，并发送该报文 2。

[67]在一个例子中，在“从起始第二区块与终止第二区块之间的每个第二区块的第一存储空间中读取报文片段”之后，还可以释放所读取的报文片段占用的第一存储空间。当某个第一区块中的每个第二区块的第一存储空间均被释放时，则可以将所述第一区块回收为空闲的第一区块，例如将这个第一区块回收到空闲地址池。然后，网络设备可以使用这个第一区块存储新的报文片段。此外，网络设备还可以根据该第一区块指向的下一个第一区块的地址，更新队列的头指针。

[68]在一个例子中，在将所述第一区块回收为空闲的第一区块后，还可以释放所述第一区块中的每个第二区块的第二存储空间。

[69]或者，在将所述第一区块回收为空闲的第一区块后，还可以不释放所述第一区块中的每个第二区块的第二存储空间，而是调整所述第一区块的反转标记。例如，网络设备在每次生成第二区块对应的 PD 时，可以获取该第二区块所属的第一区块当前的反转标记，并确定该第二区块对应的 PD 包括该反转标记。其中，第一区块连续两次被使用时的反转标记不同。例如，第一区块在第一次被使用时的反转标记为 0，在被回收为空闲

后的第二次被使用时的反转标记被修改为与 0 不同（例如可为 1），在再次被回收为空闲后的第三次被使用时的反转标记被修改为与 1 不同（例如可为 0 或 2），以此类推。

[70]例如，在第一次使用第一区块 1 存储报文时，在第一区块 1 中的每个第二区块的第二存储空间写入 PD。参见图 3A 所示，为第二存储空间写入 PD 的示例。在针对第一区块 1 的读操作完成后，如果不对这些 PD 进行处理，就会出现错误。假设，在第二次使用第一区块 1 存储报文时，只在第一区块 1 中的第二区块 0、第二区块 1、第二区块 2 中写入报文片段和 PD，而不在第一区块 1 中的其它第二区块中写入报文片段和 PD。这样，在再次针对第一区块 1 进行读操作时，只能从第二区块 0、第二区块 1、第二区块 2 读取到报文片段，但可以从第二区块 0 至第二区块 31 读取到 PD，从而使得在使用这些 PD 对读取到的报文片段进行组包时会出现错误。

[71]针对上述发现，在一个例子中，在针对第一区块 1 的读操作完成后，删除第一区块 1 中的每个第二区块的第二存储空间的 PD。但是，在该方式下，每次针对第一区块 1 的读操作完成后，均需要删除第一区块 1 的每个第二区块的第二存储空间的 PD，导致第二区块的反复操作，并降低网络设备的处理性能。

[72]在另一个例子中，为每个第一区块设置反转标记，该反转标记可以为第一数值、例如 0 或者第二数值、例如 1。基于此，每次针对第一区块的读操作完成之后，在将该第一区块回收为空闲时，若该第一区块的前次反转标记为第一数值，则可以将该第一区块的当前反转标记修改为第二数值；反之，若该第一区块的前次反转标记为第二数值，则可以将该第一区块的当前反转标记修改为第一数值。

[73]在此基础上，每次在第一区块中的第二区块的第二存储空间存储 PD 时，所述 PD 还可以包括该第一区块的当前反转标记。例如，如图 3B 和 3D 所示，第二区块的第二存储空间中的 PD 除了包括上述的信息域 LEN、SOP、EOP、BV、VAL、NBA、NBN 之外，还可包括信息域 SEL 以记录对应的第一区块的当前反转标记。或者，如图 3C 和 3E 所示，第二区块的第二存储空间中的 PD 可包括两套信息域 LEN、SOP、EOP、BV、VAL、NBA、NBN。其中，信息域 LEN0、SOP0、EOP0、BV0、VAL0、NBA0、NBN0 有效时表示对应第一区块的当前反转标记为第一数值，信息域 LEN1、SOP1、EOP1、BV1、VAL1、NBA1、NBN1 有效时表示对应第一区块的当前反转标记为第二数值。

[74]这样，若第一区块的当前反转标记为 0，针对图 3A，对应的 PD 内的信息域 SEL 可以如图 3B 所示取值为 0，图 3B 中简化示出为“SEL0”，以表示第一区块当前的反转标记为 0。或者如图 3C 所示，对应的 PD 内的信息域 SOP0、EOP0、NBN0、NBA0 有

效。信息域 SOP0、EOP0、NBN0、NBA0 中的“0”表示第一区块当前的反转标记为 0。

[75]然后，在将第一区块回收为空闲的第一区块时，由于第一区块对应的前次反转标记为 0，因此将该第一区块的当前反转标记修改为 1。再次使用该第一区块存储报文时，若只在第二区块 0、第二区块 1、第二区块 2 中写入报文和 PD，则对应的 PD 内的信息域 SEL 可以如图 3D 所示取值为 1，图 3D 中简化示出为“SEL1”，以表示第一区块当前的反转标记为 1。或者如图 3E 所示，对应的 PD 内的信息域 SOP1、EOP1 有效。信息域 SOP1、EOP1 中的“1”表示第一区块当前的反转标记为 1。

[76]通过对比图 3D 和图 3B，在针对第一区块进行读操作时，可以从第二区块 0 至第二区块 31 读取到 PD，但是只有第二区块 0、第二区块 1、第二区块 2 对应的 PD 内的信息域 SEL 的取值为 1，即只有第二区块 0、第二区块 1、第二区块 2 是在第一区块当前的反转标记 1 下被使用的，因此只使用第二区块 0、第二区块 1、第二区块 2 的第二存储空间中的 PD 对读取到的报文片段进行组包，不会发生错误。同理，通过对比图 3C 和图 3E，也可以只使用第二区块 0、第二区块 1、第二区块 2 的第二存储空间中的 PD 对读取到的报文片段进行组包。

[77]基于与上述方法同样的申请构思，本申请实施例中还提出一种报文处理装置，内存被划分成多个第一区块，每个第一区块被划分成多个第二区块，每个第二区块均包括第一存储空间和第二存储空间，如图 4 所示，所述装置包括:

[78]写入模块 401，用于获得待存储的第一报文，根据所述第一报文的长度和所述第一存储空间的长度确定所述第一报文需要占用的第二区块，并将所述第一报文存储到所述确定的第二区块的第一存储空间；针对所述确定的每个第二区块，生成该第二区块对应的报文描述符 PD，并将所述 PD 存储到该第二区块的第二存储空间；

[79]读取模块 402，用于根据待读取的第二报文的开始地址确定要读取的第二区块，从所述要读取的第二区块的第一存储空间读取报文片段，从所述要读取的第二区块的第二存储空间读取 PD，并利用所读取的 PD 将所读取的报文片段组成所述第二报文。根据一个示例，读取模块 402 还用于发送所组成的所述第二报文。需要说明的是，待存储的第一报文与待读取的第二报文可为同一报文，也可为不同的报文，本公开对此不做限定。

[80]在一个例子中，所述写入模块 401 在确定所述第一报文需要占用的第二区块的过程中，具体用于：确定所述第一报文对应的队列；若所述队列对应有尾指针，则从所述尾指针指向的第二区块开始确定所述第一报文需要占用的第二区块；若所述队列没有对应

尾指针，则选取空闲的第一区块，并从选取的第一区块中的首个第二区块开始确定所述第一报文需要占用的第二区块。

[81]在一个例子中，所述写入模块 401 在生成该第二区块对应的 PD 的过程中，具体用于：若该第二区块是所述第一报文需要占用的第一个第二区块，则在该第二区块对应的 PD 内记录所述第一报文的长度，并将该第二区块对应的 PD 内的第一标记置为有效，以表示该第二区块为存储该第一报文的第一个第二区块；若该第二区块是所述第一报文需要占用的最后一个第二区块，则将该第二区块对应的 PD 内的第二标记置为有效，以表示该第二区块为存储该第一报文的最后一个第二区块。

[82]在一个例子中，当所述第一报文需要占用的所有第二区块属于至少两个不同的第一区块时，所述写入模块 401 在生成该第二区块对应的 PD 的过程中，具体用于：若该第二区块是当前第一区块的最后一个第二区块、但不是所述第一报文需要占用的最后一个第二区块，则在该第二区块对应的 PD 内记录指向所述第一报文需要占用的下一个第一区块的指针，在该第二区块对应的 PD 内记录所述第一报文在所述下一个第一区块中需要占用的第二区块数量。

[83]在一个例子中，所述写入模块 401 在用于生成该第二区块对应的 PD 的过程中，具体用于：获取该第二区块所属的第一区块当前的反转标记，使该第二区块对应的 PD 包括所述第一区块当前的反转标记。其中，第一区块连续两次被使用时的反转标记不同。

[84]所述读取模块 402，具体用于：从所述要读取的第二区块所属的当前第一区块的每个第二区块的第二存储空间读取 PD，将第一标记有效的 PD 对应的第二区块确定为存储所述第二报文的起始第二区块，将第二标记有效的 PD 对应的第二区块确定为存储所述第二报文的终止第二区块；若从所述当前第一区块未读取到第二标记有效的 PD，则从所述当前第一区块中的最后一个第二区块的第二存储空间读取 PD，以从该 PD 获取指向下一个第一区块的指针、以及获取所述第二报文在所述下一个第一区块中占用的第二区块数量，并根据所述指针和所述第二区块数量，从所述下一个第一区块中的第二区块的第二存储空间读取 PD。以此类推，直至读取到第二标记有效的 PD，并将第二标记有效的 PD 对应的第二区块确定为存储所述第二报文的终止第二区块；从起始第二区块与终止第二区块之间的每个第二区块的第一存储空间中读取报文片段，并将所读取的报文片段组成所述第二报文，并发送所组成的所述第二报文。

[85]所述读取模块 402，还用于：在从起始第二区块与终止第二区块之间的每个第二区块的第一存储空间中读取报文片段之后，释放所读取的报文片段占用的第一存储空间；

当某个第一区块中的每个第二区块的第一存储空间均被释放时，将所述第一区块回收为空闲；在将所述第一区块回收为空闲后，释放所述第一区块中的每个第二区块的第二存储空间，或者，调整所述第一区块的当前反转标记。

[86]基于与上述方法同样的申请构思，本申请实施例中还提出一种网络设备，本申请实施例提供的网络设备，从硬件层面而言，硬件架构示意图具体参见图 5A 所示。所述网络设备包括转发芯片 510 和存储单元 520，所述存储单元 520 是与所述转发芯片 510 分开部署的，如所述存储单元 520 可以是 DDR 内存，当然所述存储单元 520 也可以是其它类型的存储单元，对此不做限制。

[87]其中，所述存储单元 520 的内存被划分成多个第一区块，每个第一区块被划分成多个第二区块，每个第二区块均包括第一存储空间和第二存储空间。

[88]所述转发芯片 510，用于获得待存储的第一报文，根据所述第一报文的长度和所述第一存储空间的长度，确定所述第一报文需要占用的第二区块，将所述第一报文存储到存储单元520中的所述确定的第二区块的第一存储空间;针对所述确定的每个第二区块，生成该第二区块对应的报文描述符 PD，并将所述 PD 存储到该第二区块的第二存储空间；以及，根据待读取的第二报文的开始地址确定要读取的第二区块，从存储单元 520 中所述要读取的第二区块的第一存储空间读取报文片段，从存储单元 520 中所述要读取的第二区块的第二存储空间读取 PD，并利用读取的 PD 将读取的报文片段组成第二报文，并发送组成的第二报文。

[89]其中，转发芯片 510 可以实现上述报文处理方法，包括执行上述报文处理方法的各个操作。例如，可以将用于实现上述报文处理方法的流程写入到转发芯片 510，使转发芯片 510 可以执行上述报文处理方法的各个操作，在此不再赘述。

[90]其中，转发芯片 510 可以包括但不限于：FPGA（Field－Programmable Gate Array，现场可编程门阵列）、CPLD（Complex Programmable Logic Device，复杂可编程逻辑器件）等，对此转发芯片 510 的类型不做限制，只要能够实现上述功能即可。

[91]在一个例子中，由于存储单元 520 是转发芯片 510 的片外存储器、例如 DDR，因此，通过将报文和 PD 存储到同一个内存，使得转发芯片 510 可以从存储单元 520 的同一个内存中读取到报文和 PD，读取效率相对很高。

[92]在另一个例子中，参见图 5B 所示，上述转发芯片 510 还可以包括转发引擎 511 和 TM（Traffic Manage，流量管理）模块 512。转发引擎 511 在接收到报文后，可以对报

文进行流分类 511-1 和流量监管（如 CAR（Committed Access Rate，承诺访问速率）限速操作等）511-2，并将报文输出给 TM 模块 512。

[93]TM 模块 512 在接收到报文后，可以对报文进行拥塞避免 512-1、拥塞管理 512-2、流量整形 512-3 等操作。其中，拥塞避免 512-1 可包括如 WRED（Weighted Random Early Detection，加权随机早期检测）、尾丢弃等；拥塞管理 512-2 可包括如 PQ（Priority Queue，优先级队列）调度、WFQ（Weighted Fair Queuing，加权公平队列）调度等。此外，TM 模块 512 在接收到报文时，可以根据报文所属的队列，将报文缓存到存储单元 520 中。TM 模块 512 在需要发送报文时，根据报文所属的队列，从存储单元 520 中读出报文并发送。

[94]而 TM 模块 512 将报文缓存到存储单元 520、TM 模块 512 从存储单元 520 中读出报文并发送的过程，也就是上述报文处理方法，在此不再重复赘述。

[95]以上所述仅为本申请的实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本申请的权利要求范围之内。

权利要求书

1. 一种报文处理方法，其特征在于，内存被划分成多个第一区块，每个所述第一区块被划分成多个第二区块，每个所述第二区块均包括第一存储空间和第二存储空间，所述方法包括:

获得待存储的第一报文;

根据所述第一报文的长度和所述第一存储空间的长度，确定所述第一报文需要占用的一个或多个第二区块;

将所述第一报文存储到所述确定的第二区块的第一存储空间;

针对所述确定的每个第二区块，生成该第二区块对应的报文描述符 PD，并将所述 PD 存储到该第二区块的第二存储空间;

根据待读取的第二报文的开始地址确定要读取的一个或多个第二区块;

从所述要读取的各第二区块的第一存储空间读取报文片段，从所述要读取的第二区块的第二存储空间读取 PD;

利用所读取的 PD 将所读取的报文片段组成所述第二报文，并发送所述第二报文。

2. 根据权利要求 1 所述的方法，其特征在于，确定所述第一报文需要占用的一个或多个第二区块的过程，具体包括:

确定所述第一报文对应的队列;

若所述队列有尾指针，则从所述尾指针指向的第二区块开始，确定所述第一报文需要占用的第二区块;

若所述队列没有尾指针，则选取空闲的第一区块，并从选取的第一区块中的首个第二区块开始，确定所述第一报文需要占用的第二区块。

3. 根据权利要求 1 所述的方法，其特征在于，生成该第二区块对应的 PD 的过程，具体包括:

若该第二区块是所述第一报文需要占用的第一个第二区块，则在该第二区块对应的 PD 内记录所述第一报文的长度，并将该第二区块对应的 PD 内的第一标记置为有效以表示该第二区块为存储所述第一报文的第一个第二区块;

若该第二区块是所述第一报文需要占用的最后一个第二区块，则将该第二区块对应的 PD 内的第二标记置为有效以表示该第二区块为存储所述第一报文的最后一个第二区块。

4. 根据权利要求 1 所述的方法，其特征在于，生成该第二区块对应的 PD，包括:

若该第二区块是当前第一区块的最后一个第二区块、但不是所述第一报文需要占用

的最后一个第二区块，则在该第二区块对应的 PD 内记录指向所述第一报文需要占用的下一个第一区块的指针，并在该第二区块对应的 PD 内记录所述第一报文在所述下一个第一区块中需要占用的第二区块数量。

5. 根据权利要求 1 所述的方法，其特征在于，生成该第二区块对应的 PD 的过程，具体包括:

获取该第二区块所属的第一区块当前的反转标记，其中，所述第一区块连续两次被使用时的反转标记不同；并

确定该第二区块对应的 PD 包括所述第一区块当前的反转标记。

6. 根据权利要求 1 所述的方法，其特征在于，从所述要读取的第二区块的第一存储空间读取报文片段，从所述要读取的第二区块的第二存储空间读取 PD，具体包括:

从所述要读取的第二区块所属的当前第一区块的每个第二区块的第二存储空间读取 PD，将第一标记有效的 PD 对应的第二区块确定为存储所述第二报文的起始第二区块;

若从所述当前第一区块读取到第二标记有效的 PD，则将第二标记有效的 PD 对应的第二区块确定为存储所述第二报文的终止第二区块;

若从所述当前第一区块未读取到第二标记有效的 PD，则从所述当前第一区块的最后一个第二区块的第二存储空间读取 PD，以从该 PD 获取指向下一个第一区块的指针、从该 PD 获取所述第二报文在所述下一个第一区块中占用的第二区块数量，根据所述指针和所述第二区块数量从所述下一个第一区块的第二区块的第二存储空间读取 PD，并将第二标记有效的 PD 对应的第二区块确定为存储所述第二报文的终止第二区块;

从所述起始第二区块与所述终止第二区块之间的每个第二区块的第一存储空间中读取报文片段。

7. 根据权利要求 6 所述的方法，其特征在于，还包括:

释放所读取的报文片段占用的第一存储空间;

当所述当前第一区块中的每个第二区块的第一存储空间均被释放时，将所述当前第一区块回收为空闲，并释放所述当前第一区块中的每个第二区块的第二存储空间或调整所述当前第一区块的反转标记。

8. 一种报文处理装置，其特征在于，内存被划分成多个第一区块，每个所述第一区块被划分成多个第二区块，每个所述第二区块均包括第一存储空间和第二存储空间，所述装置包括:

写入模块，用于获得待存储的第一报文，并根据所述第一报文的长度和所述第一存

储空间的长度，确定所述第一报文需要占用的一个或多个第二区块，并将所述第一报文存储到所述确定的第二区块的第一存储空间；针对所述确定的每个第二区块，生成该第二区块对应的报文描述符 PD，并将所述 PD 存储到该第二区块的第二存储空间；

读取模块，用于根据待读取的第二报文的开始地址确定要读取的一个或多个第二区块，从所述要读取的各第二区块的第一存储空间读取报文片段，并从所述要读取的第二区块的第二存储空间读取 PD，并利用所读取的 PD 将所读取的报文片段组成所述第二报文，并发送所组成的所述第二报文。

9. 根据权利要求 8 所述的装置，其特征在于，所述写入模块，具体用于：

确定所述第一报文对应的队列；

若所述队列有尾指针，则从所述尾指针指向的第二区块开始，确定所述第一报文需要占用的第二区块；

若所述队列没有尾指针，则选取空闲的第一区块，并从选取的第一区块中的首个第二区块开始，确定所述第一报文需要占用的第二区块。

10. 根据权利要求 8 所述的装置，其特征在于，所述写入模块，具体用于：

若该第二区块是所述第一报文需要占用的第一个第二区块，则在该第二区块对应的 PD 内记录所述第一报文的长度，并将该第二区块对应的 PD 内的第一标记置为有效以表示该第二区块为存储所述第一报文的第一个第二区块；

若该第二区块是所述第一报文需要占用的最后一个第二区块，则将该第二区块对应的 PD 内的第二标记置为有效，以表示该第二区块为存储所述第一报文的最后一个第二区块。

11. 根据权利要求 8 所述的装置，其特征在于，所述写入模块，具体用于：

若该第二区块是当前第一区块的最后一个第二区块、但不是所述第一报文需要占用的最后一个第二区块，则在该第二区块对应的 PD 内记录指向所述第一报文需要占用的下一个第一区块的指针，在该第二区块对应的 PD 内记录所述第一报文在所述下一个第一区块中需要占用的第二区块数量。

12. 根据权利要求 8 所述的装置，其特征在于，所述写入模块，具体用于：

获取该第二区块所属的第一区块当前的反转标记，其中，所述第一区块连续两次被使用时的反转标记不同；并

确定该第二区块对应的 PD 包括所述第一区块当前的反转标记。

13. 根据权利要求 8 所述的装置，其特征在于，所述读取模块，具体用于：

从所述要读取的第二区块所属的当前第一区块的每个第二区块的第二存储空间读

取 PD，将第一标记有效的 PD 对应的第二区块确定为存储所述第二报文的起始第二区块；

若从所述当前第一区块读取到第二标记有效的 PD，则将第二标记有效的 PD 对应的第二区块确定为存储所述第二报文的终止第二区块；

若从所述当前第一区块未读取到第二标记有效的 PD，则从所述当前第一区块的最后一个第二区块的第二存储空间读取 PD，以从该 PD 获取指向下一个第一区块的指针、从该 PD 获取所述第二报文在所述下一个第一区块中占用的第二区块数量，根据所述指针和所述第二区块数量从所述下一个第一区块的第二区块的第二存储空间读取 PD，并将第二标记有效的 PD 对应的第二区块确定为存储所述第二报文的终止第二区块；

从所述起始第二区块与所述终止第二区块之间的每个第二区块的第一存储空间中读取报文片段。

14. 根据权利要求 13 所述的装置，其特征在于，所述读取模块，还用于：

释放所读取的报文片段占用的第一存储空间；

当所述当前第一区块中的每个第二区块的第一存储空间均被释放时，将所述当前第一区块回收为空闲，并释放所述当前第一区块中的每个第二区块的第二存储空间或调整所述当前第一区块的反转标记。

15、一种网络设备，包括：

存储单元，所述存储单元的内存被划分成多个第一区块，每个所述第一区块被划分成多个第二区块，每个所述第二区块均包括第一存储空间和第二存储空间；

转发芯片，用于

获得待存储的第一报文，并根据所述第一报文的长度和所述第一存储空间的长度，确定所述第一报文需要占用的一个或多个第二区块，将所述第一报文存储到所述存储单元中的所述确定的第二区块的第一存储空间；针对所述确定的每个第二区块，生成该第二区块对应的报文描述符 PD，并将所述 PD 存储到该第二区块的第二存储空间；

根据待读取的第二报文的开始地址确定要读取的一个或多个第二区块，从所述存储单元中的所述要读取的各第二区块的第一存储空间读取报文片段，从所述存储单元中的所述要读取的第二区块的第二存储空间读取 PD，并利用所读取的 PD 将所读取的报文片段组成所述第二报文，并发送所组成的所述第二报文。

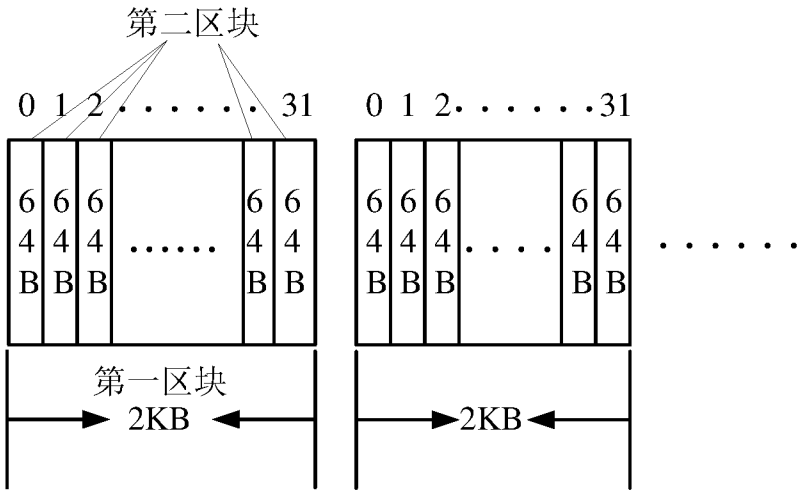

第二区块
0 1 2 31
0 1 2 31
64B
第一区块
2KB
2KB

图 1A

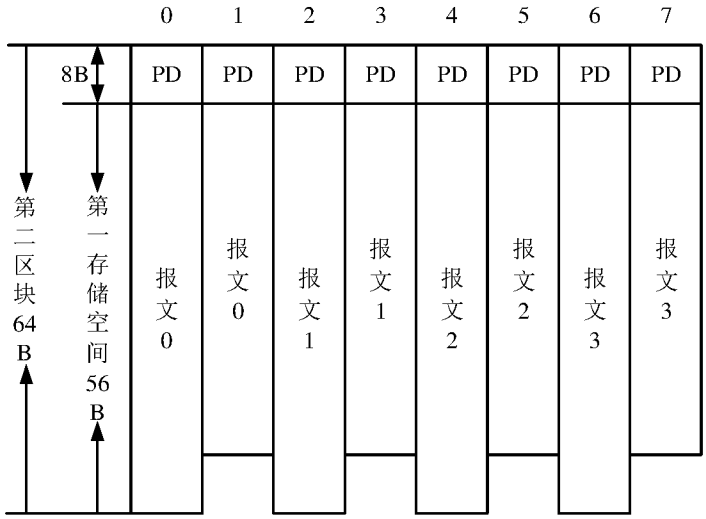

0 1 2 3 4 5 6 7
8B
PD
第二区块64B
第一存储空间56B
报文0
报文1
报文2
报文3

图 1B

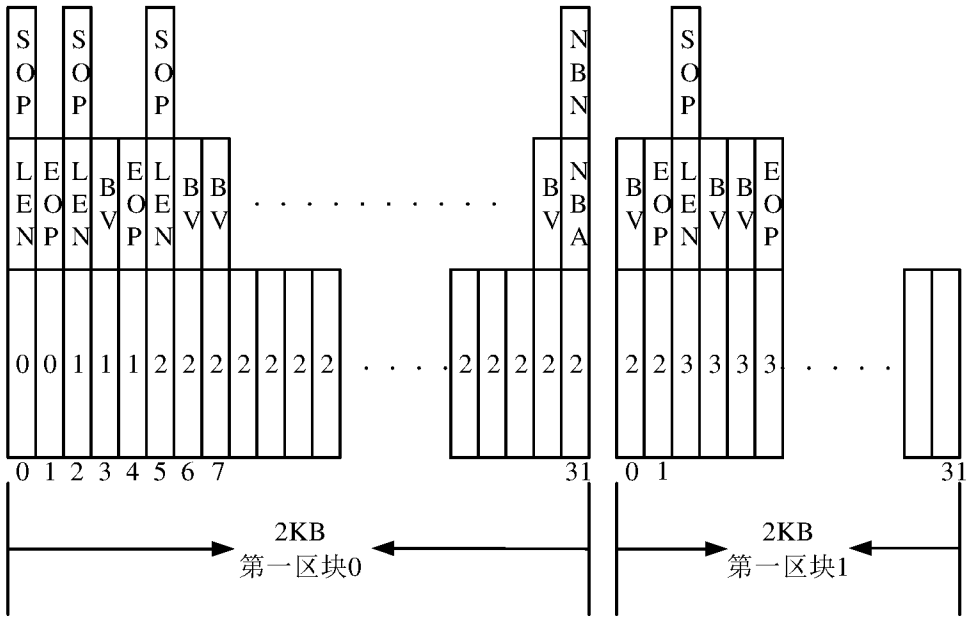

SOP
LEN
EOP
BV
NBN
NBA
0 1 2 3 4 5 6 7
31
2KB
第一区块0
第一区块1

图 1C

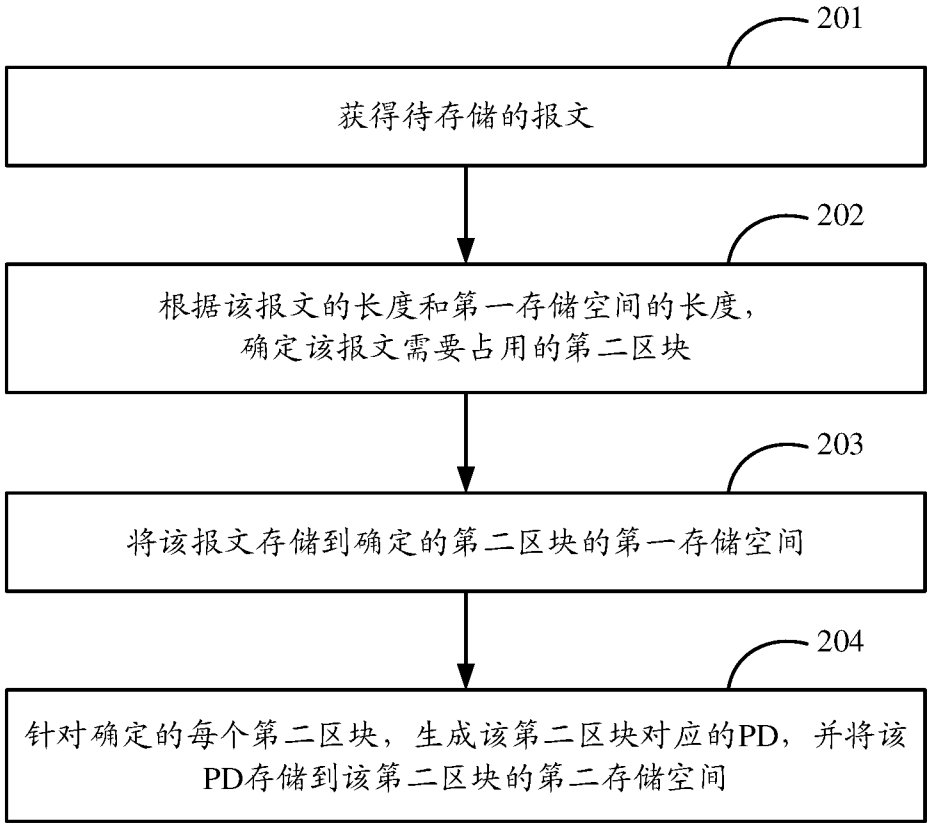

201
获得待存储的报文
202
根据该报文的长度和第一存储空间的长度，
确定该报文需要占用的第二区块
203
将该报文存储到确定的第二区块的第一存储空间
204
针对确定的每个第二区块，生成该第二区块对应的PD，并将该
PD存储到该第二区块的第二存储空间

图 2A

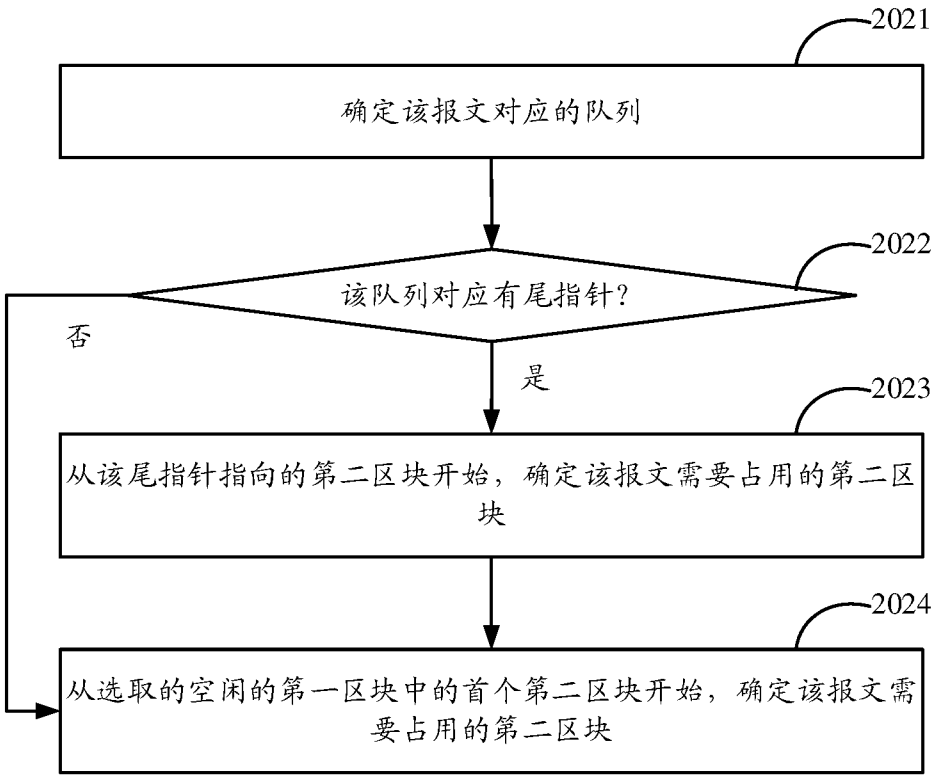

2021
确定该报文对应的队列
2022
该队列对应有尾指针？
否
是
2023
从该尾指针指向的第二区块开始，确定该报文需要占用的第二区
块
2024
从选取的空闲的第一区块中的首个第二区块开始，确定该报文需
要占用的第二区块

图 2B

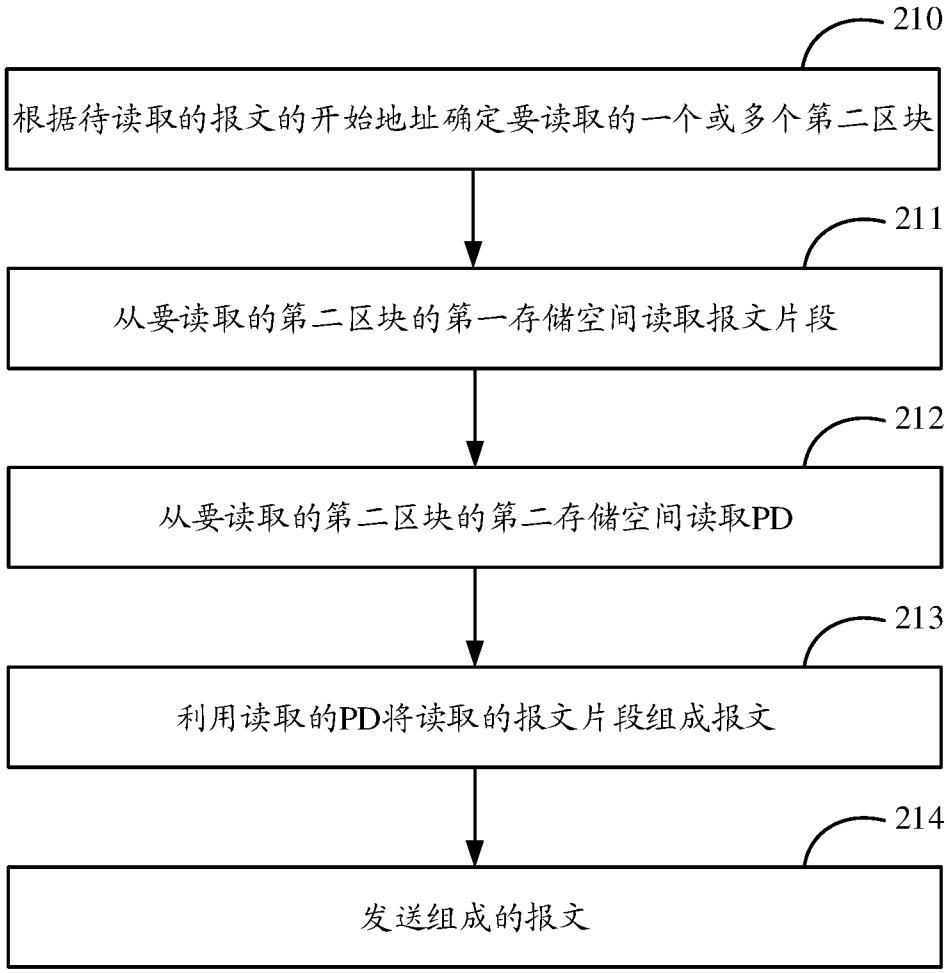

210
根据待读取的报文的开始地址确定要读取的一个或多个第二区块
211
从要读取的第二区块的第一存储空间读取报文片段
212
从要读取的第二区块的第二存储空间读取PD
213
利用读取的PD将读取的报文片段组成报文
214
发送组成的报文

图 2C

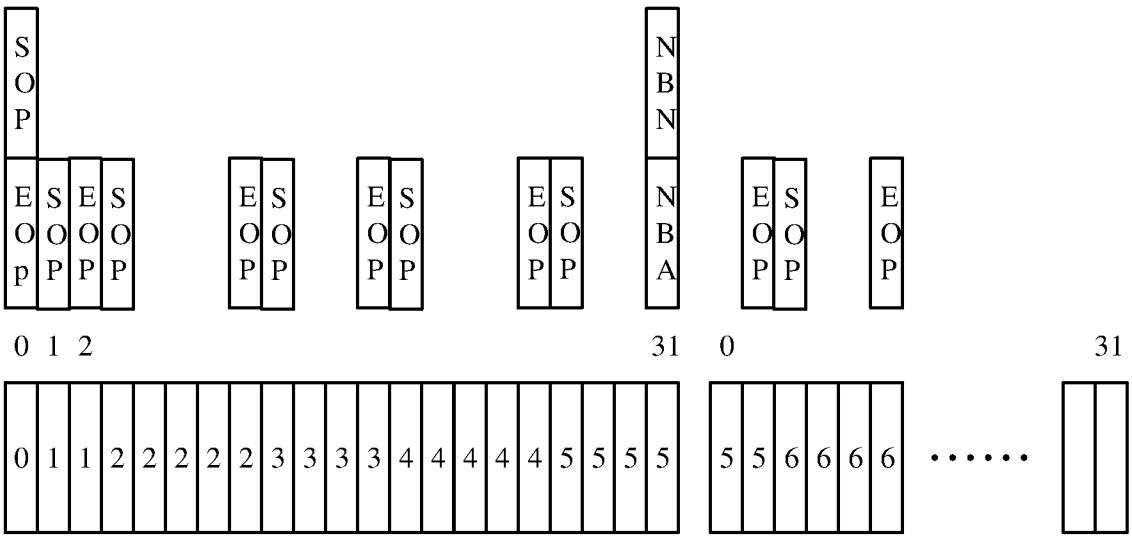

SOP
EOp
SOP
EOP
SOP
EOP
SOP
EOP
SOP
EOP
SOP
NBN
NBA
EOP
SOP
EOP
0 1 2
31
0
31
0 1 1 2 2 2 2 2 3 3 3 3 4 4 4 4 4 5 5 5 5
5 5 6 6 6 6

图 3A

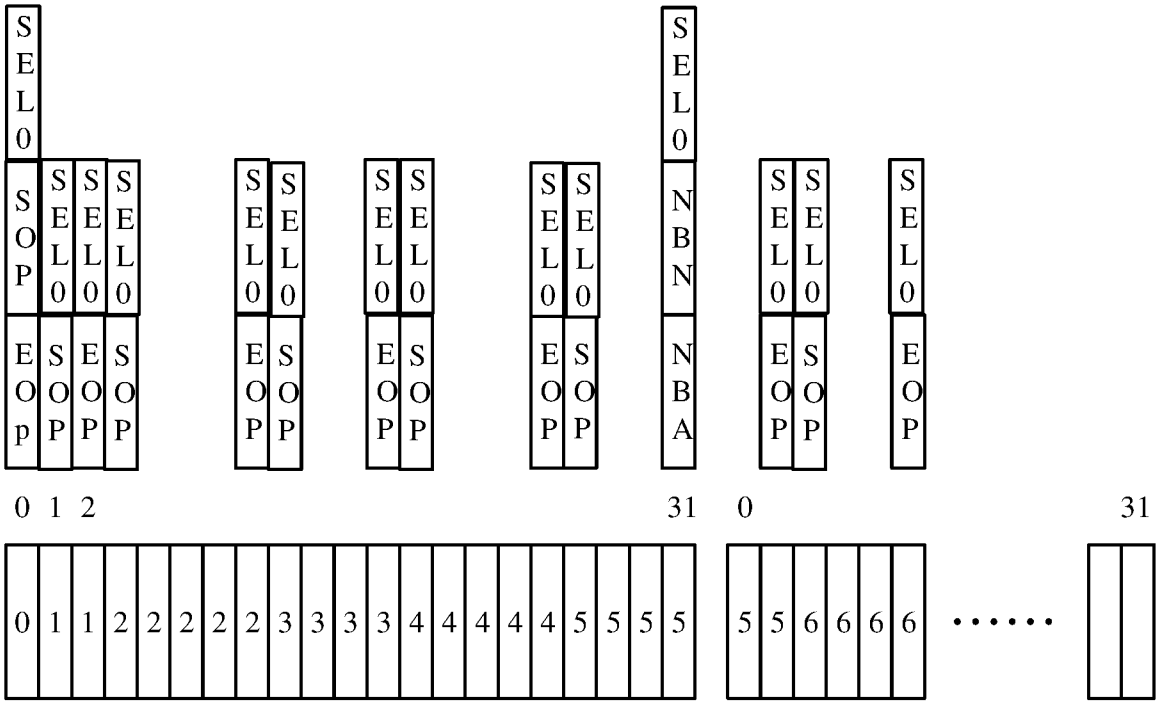

SEL0
SOP
SEL0
SEL0
SEL0
EOp
SOP
EOP
SOP
SEL0
SEL0
EOP
SOP
SEL0
SEL0
EOP
SOP
SEL0
SEL0
EOP
SOP
SEL0
NBN
NBA
SEL0
SEL0
EOP
SOP
SEL0
EOP
0 1 2
31
0
31
0 1 1 2 2 2 2 2 3 3 3 3 4 4 4 4 4 5 5 5 5
5 5 6 6 6 6
.....

图 3B

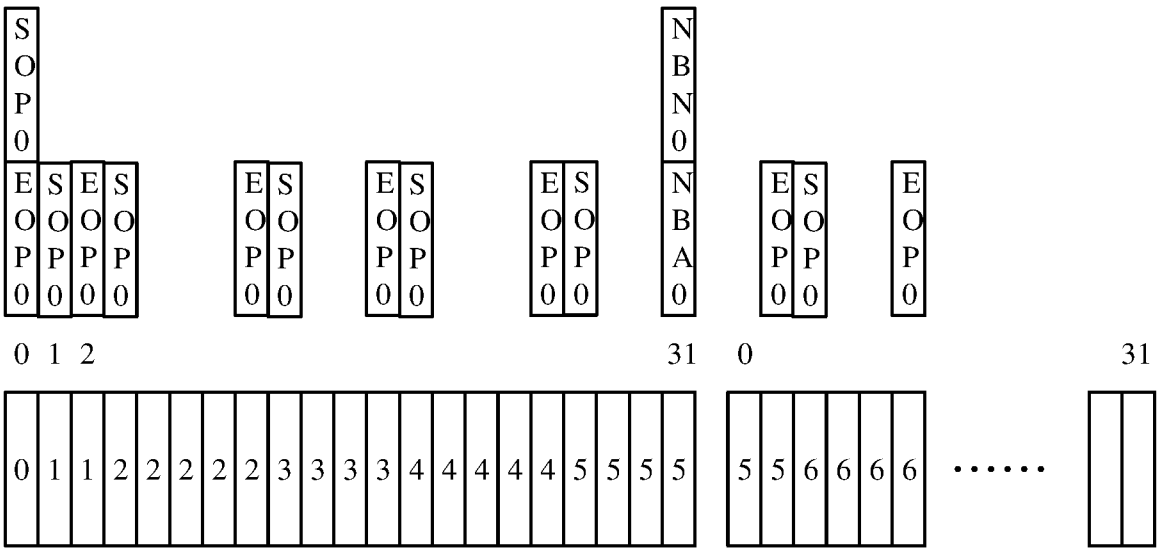

SOP0
EOP0
SOP0
EOP0
SOP0
EOP0
SOP0
EOP0
SOP0
EOP0
SOP0
NBN0
NBA0
EOP0
SOP0
EOP0
0 1 2
31
0
31
0 1 1 2 2 2 2 2 3 3 3 3 4 4 4 4 4 5 5 5 5
5 5 6 6 6 6
.....

图 3C

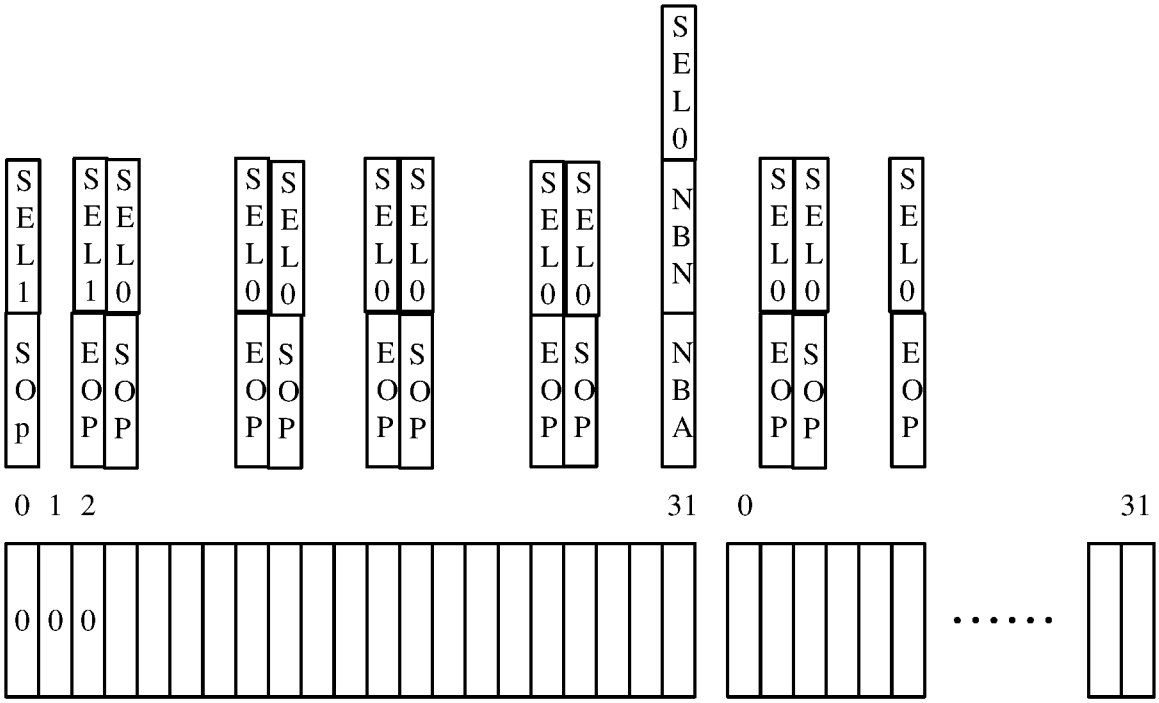

SEL1
SOp
SEL1
EOP
SEL0
SOP
SEL0
EOP
SEL0
SOP
SEL0
EOP
SEL0
SOP
SEL0
EOP
SEL0
SOP
SEL0
NBN
NBA
SEL0
EOP
SEL0
SOP
SEL0
EOP
0 1 2
31
0
31
0 0 0

图 3D

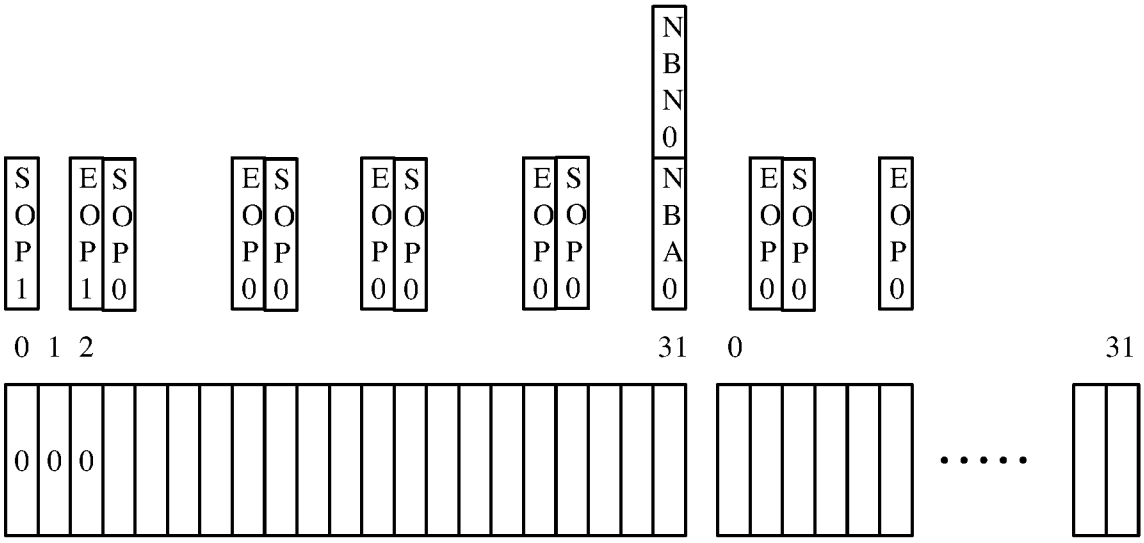

SOP1
EOP1
SOP0
EOP0
SOP0
EOP0
SOP0
EOP0
SOP0
NBN0
NBA0
EOP0
SOP0
EOP0
0 1 2
31
0
31
0 0 0

图 3E

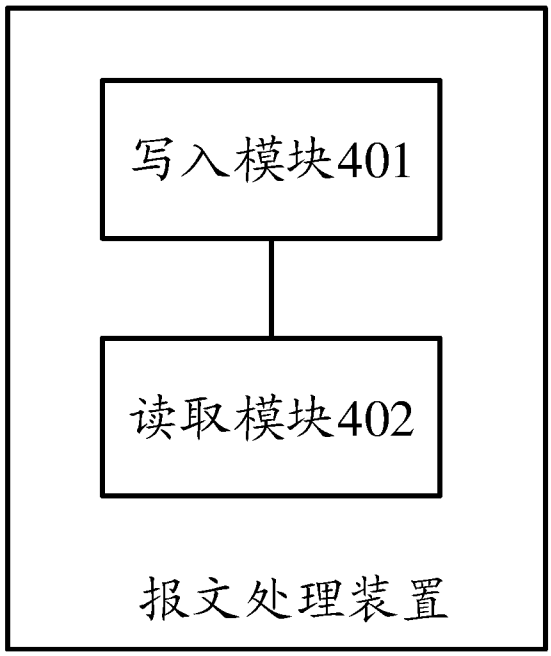


图 4

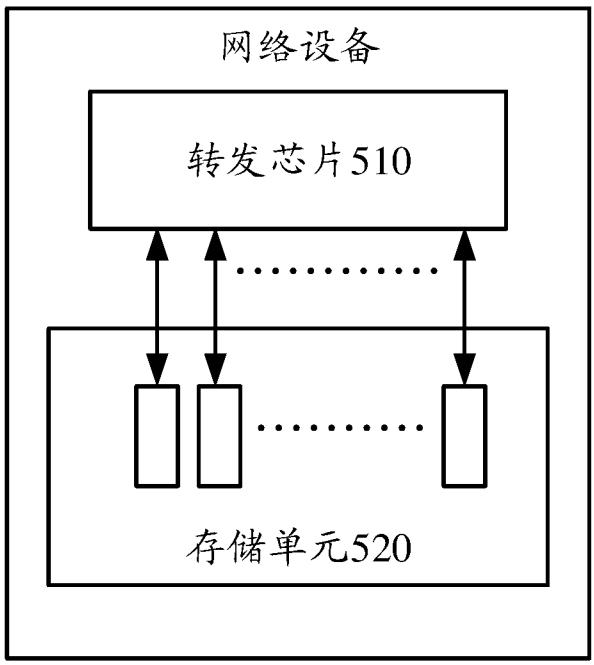


图 5A

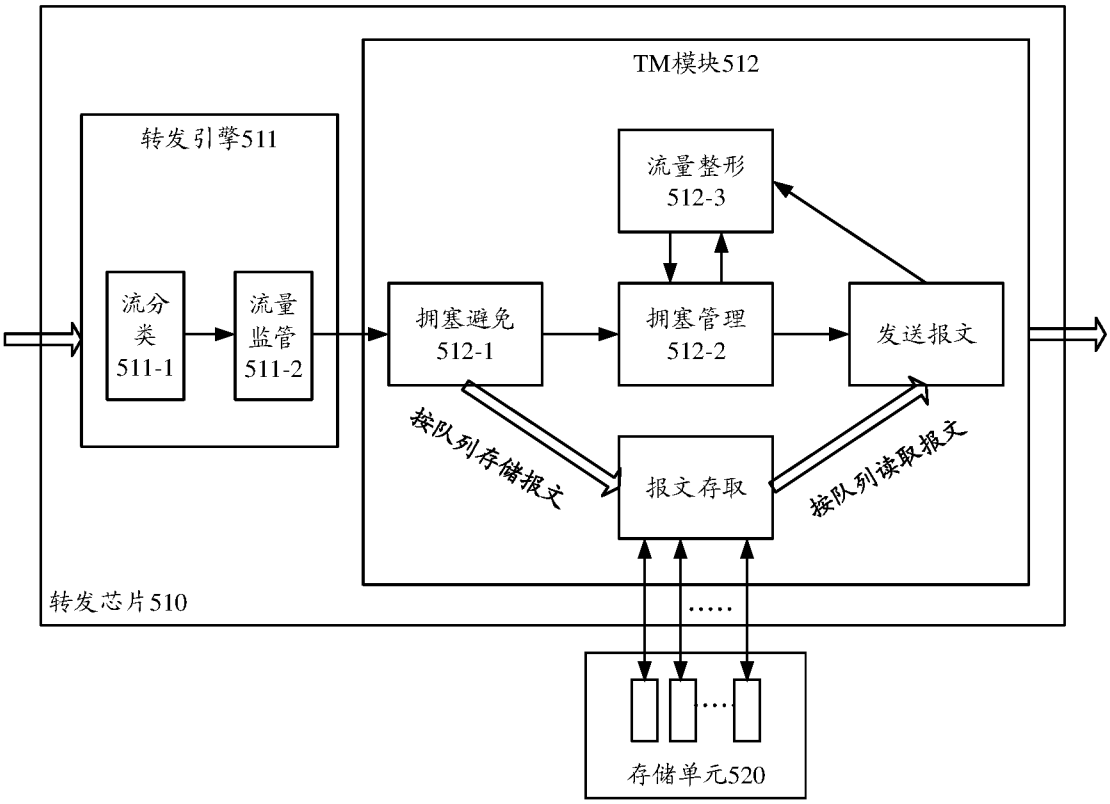

TM模块512
转发引擎511
流量整形
512-3
流分
类
511-1
流量
监管
511-2
拥塞避免
512-1
拥塞管理
512-2
发送报文
按队列存储报文
报文存取
按队列读取报文
转发芯片510
存储单元520

图 5B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/097254

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/861(2013.01)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)

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)

CPRSABS, CNTXT, VEN, CNKI, USTXT: 报文, 处理, 存储, 读, 取, 内存, 划分, 区块, 块, 片段, 长度, 描述, 符, 标志, 标识, message, storag+, read+, memory, divid+, cut+, block?, segment?, length, describ+, reqresent, sign, identif+, ID

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 101621469 A (H3C TECHNOLOGIES CO., LIMITED) 06 January 2010 (2010-01-06) entire document	1-15
A	CN 105099948 A (H3C TECHNOLOGIES CO., LIMITED) 25 November 2015 (2015-11-25) entire document	1-15
A	WO 2016078212 A1 (ZTE CORPORATION) 26 May 2016 (2016-05-26) entire document	1-15
A	US 2015095439 A1 (ALCATEL-LUCENT) 02 April 2015 (2015-04-02) entire document	1-15

☐ 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	Date of mailing of the international search report
08 October 2018	**24 October 2018**

Name and mailing address of the ISA/CN	Authorized officer
State Intellectual Property Office of the P. R. China **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing** **100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/097254

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	101621469	A	06 January 2010	CN	101621469	B	04 January 2012
				US	8225026	B2	17 July 2012
				US	2011040923	A1	17 February 2011
CN	105099948	A	25 November 2015	CN	105099948	B	15 June 2018
WO	2016078212	A1	26 May 2016	EP	3223478	A4	06 December 2017
				CN	105656804	A	08 June 2016
				JP	2017536042	A	30 November 2017
				EP	3223478	A1	27 September 2017
US	2015095439	A1	02 April 2015	US	8923204	B2	30 December 2014
				US	9246798	B2	26 January 2016
				US	2014056214	A1	27 February 2014

国际检索报告

国际申请号 PCT/CN2018/097254

A. 主题的分类

H04L 12/861(2013.01)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)

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））

CPRSABS, CNTXT, VEN, CNKI, USTXT:报文，处理，存储，读，取，内存，划分，区块，块，片段，长度，描述，符，标志，标识，message, storag+, read+, memory, divid+, cut+, block?, segment?, length, describ+, reqre-sent, sign, identif+, ID

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 101621469 A (杭州华三通信技术有限公司) 2010年 1月 6日 (2010-01-06) 全文	1-15
A	CN 105099948 A (杭州华三通信技术有限公司) 2015年 11月 25日 (2015-11-25) 全文	1-15
A	WO 2016078212 A1 (中兴通讯股份有限公司) 2016年 5月 26日 (2016-05-26) 全文	1-15
A	US 2015095439 A1 (阿尔卡特朗讯公司) 2015年 4月 2日 (2015-04-02) 全文	1-15

☐ 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	国际检索报告邮寄日期
2018年 10月 8日	2018年 10月 24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	受权官员 吕小倩 电话号码 62089460

表 PCT/ISA/210 (第2页) (2015年1月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/097254

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	101621469	A	2010年 1月 6日	CN	101621469	B	2012年 1月 4日
				US	8225026	B2	2012年 7月 17日
				US	2011040923	A1	2011年 2月 17日
CN	105099948	A	2015年 11月 25日	CN	105099948	B	2018年 6月 15日
WO	2016078212	A1	2016年 5月 26日	EP	3223478	A4	2017年 12月 6日
				CN	105656804	A	2016年 6月 8日
				JP	2017536042	A	2017年 11月 30日
				EP	3223478	A1	2017年 9月 27日
US	2015095439	A1	2015年 4月 2日	US	8923204	B2	2014年 12月 30日
				US	9246798	B2	2016年 1月 26日
				US	2014056214	A1	2014年 2月 27日