

(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
2022 年 1 月 27 日 (27.01.2022)



(10) 国际公布号
WO 2022/016889 A1

(51) 国际专利分类号:
H04L 12/801 (2013.01) **H04L 12/835** (2013.01)
H04L 12/803 (2013.01)

(21) 国际申请号: PCT/CN2021/081462

(22) 国际申请日: 2021 年 3 月 18 日 (18.03.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010713140.9 2020 年 7 月 22 日 (22.07.2020) 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: 李广 (LI, Guang); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。王小忠 (WANG, Xiaozhong); 中国广东

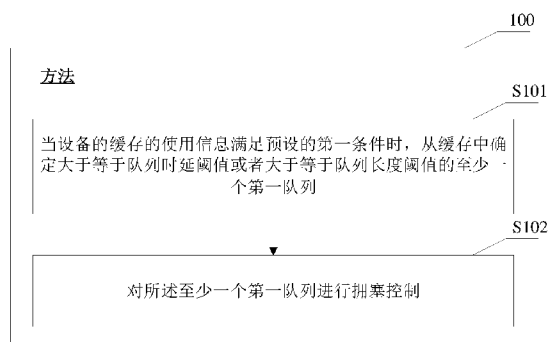
省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。袁赛 (YUAN, Sai); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。白宇 (BAI, Yu); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。张文武 (ZHANG, Wenwu); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: 深圳市深佳知识产权代理事务所 (普通合伙) (SHENPAT INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市罗湖区南湖街道春风路庐山大厦 B 座 18C2、18D、18E、18E2, Guangdong 518001 (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, IT,

(54) Title: CONGESTION CONTROL METHOD AND DEVICE

(54) 发明名称: 一种拥塞控制方法和装置



100 Method
S101 When usage information of a device cache satisfies a predetermined first condition, determine in the cache at least one first queue whose queue delay is greater than or equal to a queue delay threshold or whose queue length is greater than or equal to a queue length threshold
S102 Perform congestion control on the at least one first queue

图 2

(57) Abstract: Disclosed in the present application are a congestion control method and device, the method comprising: when usage information of a device cache satisfies a predetermined first condition, determining at least one first queue in the cache, and performing congestion control on the at least one first queue, wherein the first condition is a first threshold corresponding to the usage information of the cache, and the at least one first queue is a queue among multiple queues in the cache whose queue delay is greater than or equal to a queue delay threshold or whose queue length is greater than or equal to a queue length threshold. Thus, the device triggers a congestion control condition based on a cache state, and can achieve efficient congestion control while ensuring that the problem of buffer overflow does not occur, thereby improving the network forwarding performance of the device.

(57) 摘要: 本申请公开了一种拥塞控制方法和装置, 该方法包括: 当设备缓存的使用信息满足预设的第一条件时, 从该缓存中确定至少一个第一队列, 并对所述至少一个第一队列进行拥塞控制, 其中, 第一条件为该缓存的使用信息对应的第一阈值, 至少一个第一队列为该缓存的多个队列中队列时延大于等于队列时延阈值或者队列长度大于等于队列长度阈值的队列。这样, 设备基于缓存情况触发拥塞控制的条件, 在确保不出现缓存溢出问题的同时能够实现高效的拥塞控制, 从而提高了设备的网络转发性能。

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,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, 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,
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))。

一种拥塞控制方法和装置

本申请要求于 2020 年 07 月 22 提交中国国家知识产权局、申请号为 202010713140.9、申请名称为“一种拥塞控制方法和装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及通信技术领域，尤其涉及一种拥塞控制方法和装置。

背景技术

设备在转发过程中，如果设备接收的流量小于转发带宽，则，能够直接对所接收的流量进行转发，无需进行缓存；如果设备接收的流量大于转发带宽，则，无法对所接收的全部流量进行直接转发，需要缓存部分流量。但是，受到设备内缓存空间的限制，设备内可以缓存的流量有限，一旦所接收的流量过多，就容易出现缓存溢出的问题，为了防止设备的缓存溢出，亟待提供一种合理的拥塞控制方法，能够充分利用缓存空间，并且，避免出现误丢包或误反压的问题，从而提高设备的网络转发性能。

发明内容

基于此，本申请实施例提供了一种拥塞控制方法和装置，设备根据总体缓存空间的缓存情况设置用于触发拥塞控制的条件，并根据各队列的实际情况记录并实时更新触发拥塞控制的条件后对应待实施拥塞控制的队列，实现设备的合理拥塞控制，不仅能够充分利用缓存空间，并且能够有效的避免出现误丢包或误反压的问题，提高了设备的网络转发性能。

第一方面，本申请实施例提供了一种拥塞控制方法，该方法例如可以包括：当设备缓存的使用信息满足预设的第一条件时，从该缓存中确定至少一个第一队列，并对所述至少一个第一队列进行拥塞控制，其中，该第一条件为该缓存的使用信息对应的第一阈值，至少一个第一队列为该缓存的多个队列中队列时延大于等于队列时延阈值的队列，或者，至少一个第一队列为该缓存的多个队列中队列长度大于等于队列长度阈值的队列。这样，设备不用再基于队列的粒度进行拥塞控制，仅针对缓存空间的总体缓存情况预设用于触发拥塞控制的条件，在触发拥塞控制时设备能够针对性的对该条件对应的队列进行处理，不仅简化配置，提高了缓存空间的利用率，而且在确保不出现缓存溢出问题的同时能够实现高效的拥塞控制，从而提高了设备的网络转发性能。

其中，缓存的使用信息可以是缓存占用率或缓存剩余率，该缓存的使用信息随着设备的运行会发生变化，当有报文存入缓存或有报文从缓存中取出时，该缓存的使用信息均会发生变化。

在一种可能的实现方式中，从缓存中确定大于等于队列时延阈值的至少一个第一队列例如可以包括：确定缓存中多个队列按照队列时延从长到短排序后的前 M 个队列，作为至少一个第一队列，队列时延阈值与第 M 个队列的队列时延相等， M 为正整数。从缓存中确定大于等于队列长度阈值的至少一个第一队列例如可以包括：确定缓存中多个队列按照队

列长度从长到短排序后的前 N 个队列，作为至少一个第一队列，队列长度阈值与第 N 个队列的队列长度相等，N 为正整数。这样，当触发拥塞控制条件时，针对性的对队列长度较大或队列时延较大的队列进行控制，能够有效的提高拥塞控制的效率，使得该设备的缓存溢出的风险被快速的化解，从而提高了设备的网络转发性能。

5 作为一个示例，如果对缓存中的队列划分了优先级，那么，设备中可以针对不同的优先级设置用于触发拥塞控制的不同条件。其中，第一条件可以与第一优先级对应，至少一个第一队列属于该第一优先级。该示例中，从缓存中确定大于队列时延阈值的至少一个第一队列例如可以包括：确定第一优先级对应的多个队列按照队列时延从长到短排序后的前 K 个队列，作为至少一个第一队列，队列时延阈值与第 K 个队列的队列时延相等，K 为正整数。从缓存中确定大于队列长度阈值的至少一个第一队列例如可以包括：确定第一优先级对应的多个队列按照队列长度从长到短排序后的前 L 个队列，作为至少一个第一队列，队列长度阈值与第 L 个队列的队列长度相等，L 为正整数。这样，当触发某个优先级对应的拥塞控制条件时，针对性的对与该优先级对应的多个队列中队列长度较大或队列时延较大的队列进行控制，能够有效的提高与该优先级对应的队列的拥塞控制的效率，使得该设备的缓存溢出的风险被快速的化解，从而提高了设备的网络转发性能。

10 在一些可能的实现方式中，当缓存的使用信息为缓存占用率时，那么，第一条件可以包括缓存占用率大于第一阈值；或者，当缓存使用信息为缓存剩余率时，第一条件可以包括缓存剩余率小于第一阈值。如此，可以基于缓存空间的总体使用情况和预设的第一阈值，实现对设备中缓存的合理、有效的拥塞控制。

20 在另一些可能的实现方式中，该方法还可以包括：当缓存的使用信息的更新值满足预设的第二条件时，从缓存中确定第二条件对应的至少一个第二队列，对至少一个第二队列进行拥塞控制。其中，第二条件与第二优先级对应，至少一个第二队列属于第二优先级，第二条件为缓存的使用信息的更新值对应的第二阈值。这样，设备不用再基于队列的粒度进行拥塞控制，仅针对总体缓存空间的缓存情况设置用于触发不同拥塞控制的不同条件，能够针对性的对不同条件对应的队列进行处理，不仅简化配置，提高了缓存空间的利用率，而且确保不出现缓存溢出问题的同时能够实现高效、精准的拥塞控制，从而提高了设备的网络转发性能。

25 作为一个示例，第二优先级和第一队列所属的第一优先级不同，该情况下，针对每个优先级对应设置用于触发拥塞控制的条件，并且记录各优先级对应的被实施拥塞控制的与该优先级对应的队列，将优先级和触发拥塞控制的条件对应的关联设置，先对优先级较低的队列进行拥塞控制的方式，能够确保优先级较高的流量被设备有效的转发，提高了该设备的网络转发性能，提升了用户使用该设备的体验。

30 作为另一个示例，第二优先级和第一队列所属的第一优先级相同，该情况下，在同时使能多种拥塞控制方式且确定缓存存在溢出风险时，合理设置多种拥塞控制方式触发的先后顺序，并按照拥塞控制方式被触发的先后顺序设置对应的触发条件。这样，如果缓存中的拥塞程度较轻，则利用所确定的需要先触发的拥塞控制方式进行拥塞控制，在拥塞程度较为严重时才利用所确定的需要后触发的拥塞控制方式进行拥塞控制，将拥塞控制方式和

触发拥塞控制的条件对应的关联设置，能够确保队列的拥塞控制更加合理，提高了该设备的网络转发性能，提升了用户使用该设备的体验。

可以理解的是，缓存的使用信息的更新值为缓存占用率的更新值，第二条件包括缓存占用率的更新值大于第二阈值，第二阈值大于第一阈值；或者，缓存的使用信息的更新值为缓存剩余率的更新值，第二条件包括缓存剩余率的更新值小于第二阈值，第二阈值小于第一阈值。如此，可以基于缓存空间的总体使用情况和预设的第一阈值、第二阈值，实现对设备中缓存的合理、有效的拥塞控制。

作为一个示例，该方法中对至少一个第一队列进行拥塞控制，例如可以包括：为至少一个第一队列中的出队报文添加显示拥塞通告（英文：Explicit Congestion Notification，简称：ECN）标记，ECN 标记用于指示源端设备降低与第一队列对应的队列中的报文的发送速率。该情况下，该方法中对至少一个第二队列进行拥塞控制，例如包括：通过至少一个第二队列的入端口向设备的上游设备发送反压信号，该反压信号用于指示上游设备停止发送报文。

可以理解的是，该方法中对至少一个第一队列进行拥塞控制，例如可以包括：对至少一个第一队列进行反压控制或丢包控制。其中，反压控制是指阻止该设备的上游设备（或设备中的上游模块）发送报文或者控制源端设备降低发送报文的速率；丢包控制是指对设备接收到的报文进行丢弃。无论是反压控制还是丢包控制，均可以有效的减少缓存的存储压力，避免缓存出现溢出的现象。

丢包控制具体可以是对进入第一队列的报文在入队之前进行丢包处理，也可以是对已经进入第一队列的报文进行丢包处理。反压控制可以是向上游设备发送反压信号（例如：基于优先级流控（英文：Priority-Based Flow Control，简称：PFC）帧或停顿（英文：PAUSE）帧）实现反压，也可以是对设备内的上游模块（如上游缓存）发送反压信号实现反压。

此外，该方法中确定的至少一个第一队列还可以包括预设时长内无出队报文且无入队报文的队列，该队列称为 idle 队列，由于 idle 队列在预设时长内无入队和/或出队操作，可以确定该 idle 队列无法正常参与到设备的转发过程中，将该 idle 进行拥塞控制（例如丢包），能够释放该 idle 队列所占用的缓存空间，有效的提高缓存的利用率，以解决拥塞问题。

第二方面，本申请实施例还提供了一种拥塞控制装置包括：第一确定单元和第一控制单元。其中，第一确定单元，用于当设备缓存的使用信息满足预设的第一条件时，从缓存中确定大于等于队列时延阈值或者大于等于队列长度阈值的至少一个第一队列，第一条件为缓存的使用信息对应的第一阈值；第一控制单元，用于对至少一个第一队列进行拥塞控制。

在一些可能的实现方式中，该第一确定单元具体用于：确定缓存中多个队列按照队列时延从长到短排序后的前 M 个队列，作为至少一个第一队列，队列时延阈值与第 M 个队列的队列时延相等，M 为正整数；或者，第一确定单元具体用于：确定缓存中多个队列按照队列长度从长到短排序后的前 N 个队列，作为至少一个第一队列，队列长度阈值与第 N 个队列的队列长度相等，N 为正整数。

作为一个示例，第一条件与第一优先级对应，至少一个第一队列属于第一优先级。

该示例下，该第一确定单元，具体用于：确定第一优先级对应的多个队列按照队列时延从长到短排序后的前 K 个队列，作为至少一个第一队列，队列时延阈值与第 K 个队列的队列时延相等，K 为正整数；或者，第一确定单元，具体用于：确定第一优先级对应的多个队列按照队列长度从长到短排序后的前 L 个队列，作为至少一个第一队列，队列长度阈值与第 L 个队列的队列长度相等，L 为正整数。

可以理解的是，缓存的使用信息为缓存占用率，第一条件包括缓存占用率大于第一阈值；或者，缓存使用信息为缓存剩余率，第一条件包括缓存剩余率小于第一阈值。

在另一些可能的实现方式中，该装置还可以包括：第二确定单元和第二控制单元。其中，第二确定单元，用于当缓存的使用信息的更新值满足预设的第二条件时，从缓存中确定第二条件对应的至少一个第二队列，第二条件与第二优先级对应，至少一个第二队列属于第二优先级，第二条件为缓存的使用信息的更新值对应的第二阈值；第二控制单元，用于对至少一个第二队列进行拥塞控制。

作为一个示例，第二优先级和第一队列所属的第一优先级不同。

作为另一个示例，第二优先级和第一队列所属的第一优先级相同。

在一些可能的实现方式中，缓存的使用信息的更新值为缓存占用率的更新值，第二条件包括缓存占用率的更新值大于第二阈值，第二阈值大于第一阈值；或者，缓存的使用信息的更新值为缓存剩余率的更新值，第二条件包括缓存剩余率的更新值小于第二阈值，第二阈值小于第一阈值。

在另一些可能的实现方式中，第一控制单元具体用于：为至少一个第一队列中的出队报文添加显示拥塞通告 ECN 标记，ECN 标记用于指示源端设备降低与第一队列对应的队列中的报文的发送速率。该实现方式中，第二控制单元，具体用于：通过至少一个第二队列的入端口向设备的上游设备发送第一反压信号，第一反压信号用于指示上游设备停止发送报文。

可以理解的是，第一控制单元具体用于：对至少一个第一队列进行反压控制或丢包控制。

在一些可能的实现方式中，至少一个第一队列还包括预设时长内无出队报文且无入队报文的队列。

需要说明的是，第二方面提供的拥塞控制装置，对应于第一方面提供的拥塞控制方法，故，第二方面提供的拥塞控制装置的各种可能的实现方式以及达到的技术效果，可以参照前述第一方面提供的拥塞控制方法的介绍。

第三方面，本申请实施例还提供了一种设备，该设备包括存储器和处理器。其中，存储器，用于存储程序代码；处理器，用于运行程序代码中的指令，使得网络设备执行以上第一方面中任意一种实现方式的拥塞控制方法。

第四方面，本申请实施例还提供了一种计算机程序产品，当其在计算机上运行时，使得计算机执行前述第一方面中任意一种实现方式所述的拥塞控制方法。

第五方面，本申请实施例还提供了一种计算机可读存储介质，该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，使得该计算机执行前述第一方面中任意一种实现

方式所述的拥塞控制方法。

附图说明

下面将对实施例描述中所需要使用的附图作简单地介绍。

- 5 图 1a 为本申请实施例中一种设备 100 的结构示意图；
图 1b 为本申请实施例中设备 100 中队列的示例图；
图 2 为本申请实施例中一种拥塞控制方法 100 的流程示意图；
图 3 为本申请实施例中一种设备 200 的结构示意图；
图 4 为本申请实施例中一种丢包控制的示意图；
10 图 5 为本申请实施例中另一种拥塞控制方法 200 的流程示意图；
图 6a 为本申请实施例中一种拥塞控制过程一示例的示意图；
图 6b 为本申请实施例中一种拥塞控制过程另一示例的示意图；
图 6c 为本申请实施例中一种拥塞控制过程又一示例的示意图；
图 6d 为本申请实施例中一种拥塞控制过程再一示例的示意图；
15 图 7a 为本申请实施例中一种优先级划分方式的一示意图；
图 7b 为本申请实施例中一种优先级划分方式的另一示意图；
图 8 为本申请实施例中一种拥塞控制装置 800 的结构示意图；
图 9 为本申请实施例中一种设备 900 的结构示意图。

20 具体实施方式

下面将结合附图，对本申请实施例中的技术方案进行描述。本申请实施例描述的网络架构以及业务场景是为了更加清楚的说明本申请实施例的技术方案，并不构成对于本申请实施例提供的技术方案的限定，本领域普通技术人员可知，随着网络架构的演变和新业务场景的出现，本申请实施例提供的技术方案对于类似的技术问题，同样适用。

- 25 本申请中的“1”、“2”、“3”、“第一”、“第二”以及“第三”等序数词用于对多个对象进行区分，不用于限定多个对象的顺序。

本申请中提及的“A 和/或 B”，应该理解为包括以下情形：仅包括 A，或仅包括 B，或者同时包括 A 和 B。

- 30 设备缓存是指，对于具有转发功能的设备，设备内部通常具有缓存功能。例如在设备接收的流量大于转发带宽时，设备缓存用于存储所接收的流量中无法被直接转发的报文。而报文在缓存中以队列的形式存储。缓存中的队列包括入口队列（英文：Input Queue，简称：IQ）或出口队列（英文：Output Queue，简称：OQ）。作为一个示例，缓存中通常以 OQ 的形式进行存储，即，缓存中实际存储至少一个 OQ，每个 OQ 中记录各个该 OQ 中各报文的报文标识、描述符等报文信息；在另一种示例中，该缓存中还包括虚拟存在的至少一个 IQ，每个虚拟的 IQ 记录该 IQ 的队列长度（或时延），即，有报文进入该 IQ 时，该 IQ 的队列长度（或时延）更新。当一个报文进入设备并被存储在缓存中时，一方面，接收该报文的入端口对应 IQ 根据该增加报文刷新队列长度（或时延），例如：在该入端口对应
- 35

IQ 的队列长度中增加该报文的长度得到该 IQ 更新后的队列长度；另一方面，用于发送该报文的出端口对应 OQ 中增加该报文的相关信息，例如：在该出端口对应 OQ 中增加该报文的标识，并且在该 OQ 的队列长度上增加该报文的长度以更新该 OQ 的队列长度。

举例来说，如图 1a 所示，设备 100 中包括入端口 A、入端口 B、出端口 C 和出端口 D，该设备 100 还包括缓存 110，报文 1~7 经过该设备 100 转发，具体而言，设备 100 经过入端口 A 接收报文 1~4，经过入端口 B 接收报文 5~7，报文 1、报文 3、报文 5 和报文 7 经过出端口 C 从设备 100 转出，报文 2、报文 4 和报文 6 经过出端口 D 从设备 100 转出。将报文 1~7 存储在设备 100 的缓存 110 中，如图 1b 所示，该设备 100 中存在 4 个队列，分别为：IQA、IQB、OQC 和 OQD，其中，IQA 和 IQB 可以属于虚拟存在的队列，IQA 对应存储有报文 1、报文 2、报文 3 和报文 4 的总长度（或总时延），记作队列长度 1（或时延 1）；IQB 中对应存储有报文 5、报文 6 和报文 7 的总长度（或总时延），记作队列长度 2（或时延 2）；OQC 中包括报文 1、报文 3、报文 5 和报文 7，OQD 中包括报文 2、报文 4 和报文 6，其中，OQC 和 OQD 记录其队列中报文对应的报文标识和描述符等报文信息。

将无法直接转发的报文以队列的形式存储在设备的缓存中，一定程度上解决了所接收的流量超过设备的转发带宽的情况下流量无法直接进行转发的问题。但是，由于设备上的缓存空间有限，设备上很可能出现缓存溢出的现象，即，缓存的流量过多导致缓存空间不足，此时，需要对设备进行拥塞控制。

目前，通常在队列的粒度上进行拥塞控制。

第一种方式中，设备中的每个队列单独设置对应的队列长度阈值（或队列时延阈值），当队列长度（或队列时延）超过该队列对应的队列长度阈值（或队列时延阈值）时，对该队列进行丢包或反压等方式的拥塞控制。该方式虽然可以避免设备上出现缓存溢出的现象，但是，由于各队列的队列长度（或队列时延）的不确定性，各队列的队列长度（或队列时延）往往分布十分不均匀，所以，该方式中给每个队列单独设置合理的队列长度阈值（或队列时延阈值）十分困难，如果设置的队列长度阈值（或队列时延阈值）太小，将导致流量占用了较小的缓存空间时该队列就开始进行拥塞控制，使得设备无法充分利用缓存空间，从而端口无法维持稳定的报文传输速率，使得网络吞吐量降低；如果设置的队列长度阈值（或队列时延阈值）太大，将导致缓存溢出从而使得设备进行无差别的丢包或反压等拥塞控制，使得网络转发性能降低。

第二种方式中，为了提高对设备的缓存空间的利用率，也可以为设备中的每个队列单独设置对应的阈值系数，基于设备中当前剩余的缓存空间大小和每个队列的阈值系数确定该队列对应的队列长度阈值，例如：队列的长度阈值=该队列的阈值系数 $a \times$ 当前剩余的缓存空间大小，当队列长度超过该队列对应的队列长度阈值时，对该队列进行丢包或反压等方式的拥塞控制。该方式由于引入当前剩余缓存空间的大小动态确定每个队列的队列长度阈值，所以，能够克服第一种方式中无法充分利用缓存空间的问题，但是，由于各队列的队列长度的不确定性，所以，给每个队列单独设置合理的阈值系数仍然十分困难，如果设置的阈值系数不合理，将导致设备进行误丢包或误反压等拥塞控制，使得网络转发性能降低。

基于此，本申请实施例提供了一种拥塞控制方法，该方法中不再基于队列的粒度进行拥塞控制，而是基于设备缓存的状态预设触发拥塞控制的条件，并且，基于缓存中各个队列的实时情况确定进行拥塞控制时待处理的至少一个队列，这样，当确定设备的当前缓存使用信息（如：缓存剩余率或缓存占用率）满足预设的条件时，即可获取该条件对应的待处理的至少一个队列的标识，从而根据所获取的至少一个队列的标识对所述至少一个队列进行拥塞控制，例如，对该至少一个队列进行反压或丢包处理。这样，设备无需为每个队列设置对应的队列长度阈值（或队列时延阈值），仅针对总体缓存空间的缓存情况设置用于触发拥塞控制的条件，并基于队列的实时情况为触发拥塞控制的条件记录对应的待处理队列，以使得设备在触发拥塞控制时能够针对性的对该条件对应的队列进行处理，不仅简化配置，提高了缓存空间的利用率，而且在确保不出现缓存溢出问题的同时能够实现高效的拥塞控制，从而提高了设备的网络转发性能。

举例来说：设备的缓存使用信息为缓存占用率，设置阈值 1 为 70%，且记录队列长度最长的 2 个队列为待处理队列。该场景中，如果第一时刻设备获取的缓存占用率为 50%，则，确定该第一时刻的缓存占用率小于阈值 1，不存在缓存溢出的风险，无需进行拥塞控制。如果第二时刻设备获取的缓存占用率为 75%，则，确定该第二时刻的缓存占用率超过了阈值 1，存在缓存溢出的风险，此时，触发该设备实施拥塞控制。第二时刻，假设设备的缓存中包括队列 1~队列 3，其中，队列 1 的队列长度为 50 兆（英文：MB），队列 2 的队列长度为 10MB，队列 3 的队列长度为 80 MB，则设备可以对记录待处理队列（即队列 1 和队列 3）进行拥塞控制，以提高设备的转发性能和拥塞控制效率。

可以理解的是，上述场景仅是本申请实施例提供的场景示例，本申请实施例并不限于此场景。

需要说明的是，本申请实施例中的设备可以是包括报文转发功能和/或报文缓存功能的任意设备，该设备包括但不限于交换机、路由器等网络设备。

需要说明的是，本申请实施例中的缓存的使用信息，会随着设备的运行（如设备接收报文和/或设备发送报文）实时发生变化。缓存的使用信息可以是设备中缓存空间的当前缓存占用率，也可以是设备中缓存空间的当前缓存剩余率。当缓存的使用信息为缓存占用率时，有新的报文存储到该设备的缓存中时该缓存占用率会增加，缓存占用率增加的部分为该报文的长度除以整个缓存空间的大小；或者，缓存中有报文被设备转出时该缓存占用率减少，缓存占用率减少的部分为该报文的长度除以整个缓存空间的大小。当缓存的使用信息为缓存剩余率时，有新的报文存储到该设备的缓存中时该缓存剩余率会减少，缓存剩余率减少的部分为该报文的长度除以整个缓存空间的大小；或者，缓存中有报文被设备转出时该缓存剩余率增加，缓存剩余率增加的部分为该报文的长度除以整个缓存空间的大小。一个缓存在一个时刻，其缓存占用率和缓存剩余率之和为 1。

下面结合附图，通过实施例来详细说明本申请实施例中的拥塞控制方法的具体实现方式。

图 2 为本申请实施例中一种拥塞控制方法 100 的流程示意图。该方法 100 中的针对多个队列共享一个缓存的场景，通过对多个队列中的部分或全部队列的拥塞控制避免该共享

的缓存的溢出问题。作为一个示例，方法 100 中的执行主体可以是图 1a 所示的设备 100，该方法 100 中涉及的缓存为图 1a 中设备 100 所包括的缓存 110。作为另一个示例，该方法 100 中的执行主体例如也可以是图 3 所示的设备 200，该设备 200 包括缓存 201、缓存 202、...、缓存 20N（N 为大于 1 的整数），每个缓存被多个队列共享，每个缓存对该缓存上存储的部分或全部队列进行拥塞控制避免该缓存的溢出问题，即，该方法 100 中涉及的缓存可以为图 3 中设备 200 所包括的任意一个缓存，而拥塞控制的对象即为该方法 100 所涉及的缓存上存储的队列中的全部或部分队列。

具体实现时，参见图 2，该方法 100 例如可以包括下述 S101~S102：

S101，当设备缓存的使用信息满足预设的第一条件时，从缓存中确定大于等于队列时延阈值或者大于等于队列长度阈值的至少一个第一队列。

其中，缓存的使用信息随着设备的运行过程中缓存中报文的数量的变化而变化，用于表征该缓存当前被使用的情况。作为一个示例，该缓存的使用信息可以是缓存占用值或缓存占用率，缓存占用值是指该缓存中被占用的空间大小，例如，缓存占用值为 100MB；缓存占用率是指该缓存中被占用的空间大小占该缓存的总空间大小的比例，例如，假设缓存的总空间大小为 500MB，当前被占用的空间大小为 100MB，那么，当前的缓存占用率为 $(100\text{MB}/500\text{MB})=20\%$ 。作为另一个示例，该缓存的使用信息也可以是缓存剩余值或缓存剩余率，缓存剩余值是指该缓存中未被占用的空间大小，例如，缓存剩余值为 400MB；缓存剩余率是指该缓存中未被占用的空间大小占该缓存的总空间大小的比例，例如，假设缓存的总空间大小为 500MB，当前未被占用的空间大小为 400MB，那么，当前的缓存占用率为 $(400\text{MB}/500\text{MB})=80\%$ 。需要说明的是，同一时刻，缓存占用值和缓存剩余值之和为该缓存的总空间大小，缓存占用率和缓存剩余率之和为 1。

需要说明的是，当设备获取到的缓存的使用信息为缓存占用值或缓存剩余值时，一种情况下，可以基于该缓存已知的总空间大小和所获取的缓存占用值或缓存剩余值，计算对应的缓存占用率或缓存剩余率，从而基于缓存占用率或缓存剩余率判断该设备的缓存使用信息是否满足预设的第一条件；另一种情况下，也可以直接基于缓存占用值或缓存剩余值判断该设备的缓存的使用信息是否满足预设的第一条件。当设备获取到的缓存的使用信息为缓存占用率或缓存剩余率时，可以直接基于缓存占用率或缓存剩余率判断该设备的缓存使用信息是否满足预设的第一条件。

第一条件，是指用于鉴定该设备是否存在缓存溢出风险或是否需要采取拥塞控制措施的条件。当设备当前的缓存使用信息满足了预设的第一条件，则表征设备存在缓存溢出的风险，触发该设备实施本申请实施例提供的拥塞控制方法。该第一条件可以是用户或者网管通过命令行或者配置文件等方式在设备上预先设定的。作为一个示例，该第一条件可以是缓存占用率阈值或缓存剩余率阈值，那么，S101 中设备的缓存使用信息满足预设的第一条件，具体可以指缓存占用率大于缓存占用率阈值或缓存剩余率小于缓存剩余率阈值。作为另一个示例，该第一条件也可以是缓存占用值阈值或缓存剩余值阈值，那么，S101 中设备的缓存使用信息满足预设的第一条件，具体可以指缓存占用值大于缓存占用值阈值或缓存剩余值小于缓存剩余值阈值。

需要说明的是，为了提高拥塞控制的效果和效率，设备在进行拥塞控制时，可以选择
5 对时延较大或队列长度较长的队列进行拥塞控制。那么，本申请实施例，设备还可以记
录并实时更新待被实施拥塞控制的队列（即，时延较大或队列长度较长的队列）的相关信
息，例如，设备中记录至少一个第一队列的标识，记录的内容还可以包括各第一队列对应
10 的队列时延或队列长度。其中，队列时延是指以转发带宽转发队列中所有报文所需的时
长，队列长度是指队列中包括的所有报文长度的和。当设备的缓存的使用信息满足预设
的第一条件时，设备中当前记录的待被实施拥塞控制的队列记作第一队列。

如果缓存中不对队列做优先级的划分，则在确定缓存的使用信息满足第一条件时，该
设备即可获取该缓存保存的队列中至少一个第一队列的标识，并基于该至少一个第一队列
15 的标识对至少一个第一队列进行拥塞控制，具体实现方式参见下述 S102 的相关描述。

作为一个示例，如果基于队列时延确定第一队列，则，该至少一个第一队列可以包括
缓存保存的多个队列中队列时延大于等于队列时延阈值（例如：10 毫秒）的队列。例如，
将缓存中各个队列的队列时延和预设的队列时延阈值进行比较，大于预设的队列时延阈值的
队列记作第一队列。又例如，将缓存中各个队列的队列时延从长到短排序后，前 M（M
15 为正整数，例如：M=5）个队列记作第一队列，其中，第 M 个队列的队列时延等于队列时延
阈值。需要说明的是，如果设备支持时延测量功能，则，可以直接获取到各个队列的队列
时延；如果该设备不支持时延测量功能，则，可以基于相关指标计算各个队列的队列时
延，例如，基于队列长度和转发带宽计算队列的时延，具体可以是：队列的时延=队列长
20 度÷转发带宽，其中，如果该队列对应的端口未划分优先级，则，转发带宽可以是该队列
对应的端口的总带宽；如果该队列对应的端口划分了不同的优先级，则，转发带宽可以
是该端口的总带宽和为该队列对应优先级分配的调度权重的乘积。

作为另一个示例，如果基于队列长度确定第一队列，则，该至少一个第一队列可以包
括缓存保存的多个队列中队列长度大于等于队列长度阈值（例如：300MB）的队列。例如，
将缓存中各个队列的队列长度和预设的队列长度阈值进行比较，大于预设的队列长度阈值的
25 队列记作第一队列。又例如，将缓存中各个队列的队列长度从长到短排序后，前 N（N
为正整数，例如：N=3）个队列记作第一队列，其中，第 N 个队列的队列长度等于队列长度
阈值。其中，队列长度是缓存中记录的参数之一，设备可以直接获取到各个队列的队列长
度。

例如：假设在 S101 执行的时刻，设备的缓存中包括未被划分优先级的队列 1~队列 5，
30 其中，队列 1~队列 5 的队列时延分别为：5 毫秒、10 毫秒、25 毫秒、8 毫秒和 22 毫秒，
预设个数 M 为 2，那么，设备中记录的待被实施拥塞控制的第一队列为队列 3 和队列 5。
此时，设备确定当前缓存的使用信息满足预设的第一条件，获取队列 3 的标识和队列 5 的
标识，从而该设备可以执行 S102，基于队列 3 的标识和队列 5 的标识分别对队列 3 和队列
5 进行拥塞控制。

如果缓存中对队列做优先级的划分，则设备对于不同的优先级设置不同的条件，并且
为每个优先级记录该优先级对应的待实施要拥塞控制的队列。那么，S101 中的第一条件可
35 以与第一优先级对应。在确定缓存的使用信息满足第一条件时，该设备可以确定该第一条

件对应的第一优先级需要被实施拥塞控制，此时，获取该缓存中与第一优先级对应的至少一个第一队列的标识，并基于该至少一个第一队列的标识对至少一个第一队列进行拥塞控制，具体实现方式参见下述 S102 的相关描述。

作为一个示例，如果基于队列时延确定第一队列，则，该至少一个第一队列可以包括缓存保存的与第一优先级对应的多个队列中队列时延大于等于队列时延阈值（例如：10 毫秒）的队列。例如，将缓存中与第一优先级对应的各个队列的队列时延和预设的队列时延阈值进行比较，大于预设的队列时延阈值的队列记作第一队列。又例如，将缓存中与第一优先级对应的各个队列的队列时延从长到短排序后，前 K（K 为正整数，例如：K=5）个队列记作第一队列，其中，第 K 个队列的队列时延等于队列时延阈值。

作为另一个示例，如果基于队列长度确定第一队列，则，该至少一个第一队列可以包括缓存保存的与第一优先级对应的多个队列中队列长度大于等于队列长度阈值（例如：300MB）的队列。例如，将缓存中与第一优先级对应的各个队列的队列长度和预设的队列长度阈值进行比较，大于预设的队列长度阈值的队列记作第一队列。又例如，将缓存中与第一优先级对应的各个队列的队列长度从长到短排序后，前 L（L 为正整数，例如：L=3）个队列记作第一队列，其中，第 L 个队列的队列长度等于队列长度阈值。

例如：假设在 S101 执行的时刻，设备的缓存中的队列的优先级被划分到优先级 1 和优先级 2，优先级 1 对应的队列包括队列 1~队列 5，其中，队列 1~队列 5 的队列长度分别为：100MB、50MB、300MB、250MB 和 220MB，预设长度为 200MB，那么，设备中记录的与优先级 1 对应的待被实施拥塞控制的第一队列包括队列 3、队列 4 和队列 5。此时，设备确定当前的缓存的使用信息满足与优先级 1 对应的第一条条件，则，获取与优先级 1 对应的队列 3 的标识、队列 4 的标识和队列 5 的标识，从而该设备可以执行 S102，基于队列 3 的标识、队列 4 的标识和队列 5 的标识分别对队列 3、队列 4 和队列 5 进行拥塞控制。

需要说明的是，随着队列中报文的入队和出队，队列时延和队列长度不断变化，设备内记录的待被实施拥塞控制的队列也会不断更新。当有报文入队或出队时，设备更新该队列的队列时延和/或队列长度，并基于更新后的队列时延和/或队列长度更新待被实施拥塞控制的队列。例如：假设预设个数为 2，设备的缓存中包括队列 1~队列 5，在第一时刻，队列 1~队列 5 的队列长度分别为：100MB、50MB、300MB、250MB 和 220MB，此时，设备中记录的待被实施拥塞控制的第一队列包括队列 3 和队列 4。第二时刻，队列 1 中有 10MB 的报文 1 入队，队列 4 中有 80MB 的报文 2 出队，那么，队列 1~队列 5 的队列长度分别为：110MB、50MB、300MB、170MB 和 220MB，此时，设备中记录的待被实施拥塞控制的第一队列包括队列 3 和队列 5。

在另一些可能的实现方式中，对于预设时长内无入队和/或出队操作的队列，则，可以确定该队列无法正常参与到设备的转发过程中，该队列被记作 idle 队列。为了有效的提高缓存的利用率，可以释放该 idle 队列所占用的缓存空间以解决拥塞问题。具体实现时，设备可以将该 idle 队列的队列长度和/或队列时延的记录值修改为极大值，这样，设备所记录的待被实施拥塞控制的队列中包括该 idle 队列，即，S101 中的与第一条条件对应的至少一个第一队列还可以包括预设时长（如 10 秒）内无出队报文且无入队报文的 idle 队列，从而

设备执行该方法 100 即可释放该 idle 队列所占用的缓存空间。需要说明的是，该实现方式中，修改的 idle 队列的队列时延或队列长度的记录值，而不影响该缓存使用信息，S101 中用于判断是否满足第一条件的缓存的使用信息体现缓存中的实际使用情况，取决于各队列的实际队列时延或队列长度，而与各队列的记录值无关。

5 作为一个示例，设备可以为设置一个标志位和一个定时器以识别 idle 队列，定时器的设定的周期取用于鉴定 idle 队列的预设时长，标志位的值用于指示该定时器设定的周期内是否有报文入队或出队。当队列中有入队或出队操作时，该标志位被置 1 且定时器开始计时，在定时器到达设定的周期之前，如果该队列发生入队或出队操作，则，将该标志位置 1 且定时器重启开始计时；在定时器到达设定的周期时，如果该标志位为 1，则将该标志位置 0 且定时器重启开始计时；在定时器到达设定的周期时，如果该标志位为 0，则说明该设定的周期内未发生出队或入队的操作，将该队列确定为 idle 队列。

可见，通过上述 S101 中确定设备的缓存当前的使用信息满足第一条件，触发了拥塞控制的执行，并且，通过上述 S101 中获取待实施拥塞控制的至少一个第一队列，确定了实施拥塞控制的对象，为实施 S102 中的拥塞控制提供了数据基础。

15 S102，对至少一个第一队列进行拥塞控制。

具体实现时，S102 例如可以包括：设备根据至少一个第一队列的标识确定所述至少一个第一队列，从而，该设备对至少一个第一队列中的每个第一队列进行拥塞控制。

拥塞控制的方式可以包括反压和/或丢包。一种情况下，S102 具体可以包括：对所述至少一个第一队列进行反压控制。另一种情况下，S102 具体也可以包括：对所述至少一个第一队列进行丢包控制。再一种情况下，S102 具体还可以包括：对所述至少一个第一队列进行丢包控制和反压控制，同时对确定的至少一个第一队列进行丢包控制和反压控制，能够有效的提高拥塞控制的效率。其中，反压是指阻止该设备的上游设备（或设备中的上游模块）发送报文或者控制源端设备降低发送报文的速率；丢包是指对设备接收到的报文进行丢弃。无论是反压还是丢包，均可以有效的减少缓存的存储压力，避免缓存出现溢出的现象。

25 作为一个示例，反压控制的实现方式可以是控制源端设备降低报文的发送速率，例如：在第一队列的出队报文中添加显示拥塞通告（英文：Explicit Congestion Notification，简称：ECN）标记，这样，该报文到达目的端设备时，该目的端设备可以向发送该报文的源端设备发送回复报文（例如 ACK 报文），该回复报文中携带 ECN 标记，用于告知源端设备报文转发路径中有设备出现拥塞，此时，源端设备可以基于该携带 ECN 标记的回复报文降低与第一队列对应的队列中报文的发送速率，从而达到缓解该设备中缓存溢出压力的问题。需要说明的是，该方式适用于确定的第一队列为 OQ 的情况。

30 作为另一个示例，反压控制的实现方式可以是阻止该设备的上游设备发送报文，具体实现时，设备生成反压信号，并且通过该第一队列的入端口向该设备的上游设备发送该反压信号，用于告知上游设备停止发送报文，此时，该上游设备可以基于该反压信号停止通过该第一队列的入端口向该设备发送报文，从而达到缓解该设备中缓存溢出压力的问题。例如：对于区分不同队列优先级的情况，该反压信号具体可以是基于优先级流控（英文：

Priority-Based Flow Control, 简称: PFC) 帧, 即, 设备通过该第一队列的入端口向该设备的上游设备发送该 PFC 帧, 该 PFC 帧携带该待拥塞控制的队列所属的队列优先级的标识, 用于告知上游设备停止向该设备发送该队列优先级的报文。又例如: 对于不区分队列优先级的情况, 该反压信号具体可以是停顿 (英文: PAUSE) 帧, 即, 设备通过该第一队列的入端口向该设备的上游设备发送该 PAUSE 帧, 该 PAUSE 帧用于告知上游设备停止向该设备发送报文。需要说明的是, 该方式适用于确定的第一队列为 IQ 的情况。

需要说明的是, 如果该方法 100 应用于包括多个缓存的设备 (如图 3 所示的设备 300), 则, 对于多个缓存中直接与上游设备连接的缓存, 可以生成并通过该第一队列的入端口向该设备的上游设备发送反压信号, 实现对该第一队列的反压控制。对于多个缓存中不直接与上游设备连接的缓存, 该存在溢出风险的缓存也可以生成并通过该第一队列的入端口向该缓存的上游模块发送反压信号, 这样, 上游模块接收到反压信号之后将不再通过该第一队列的入端口向该存在溢出风险的缓存发送报文, 实现对该第一队列的反压控制。在设备内部的不同模块之间发送的反压信号的格式不作具体限定。需要说明的是, 该上游模块例如可以是设备中的其他缓存, 在本申请实施例中不作具体限定。

作为一个示例, 丢包控制的实现方式可以是对待进入第一队列的报文在入队之前进行丢包处理。例如: 对待进入第一队列的报文采用尾丢包 (英文: Tail Drop, 简称: TD) 的方式直接丢弃。又例如: 对待进入第一队列的报文采用加权随机早期检测 (英文: Weighted Random Early Detection, 简称: WRED) 的方式进行丢包处理, 其中, S101 中缓存使用信息满足第一条件, 该第一条件包括两个阈值 a 和 b , 其中, $a > b$, 当第一队列的队列长度小于 b 时, 不丢弃待进入第一队列的报文; 当第一队列的队列长度大于或等于 b 且小于或等于 a 时, 随机丢弃待进入第一队列的报文, 丢弃报文的概率随着该第一队列的队列长度的增加而变大; 当第一队列的队列长度大于 a 时, 丢弃待进入第一队列的所有报文。需要说明的是, 该方式适用于确定的第一队列为 OQ 的情况。

作为另一个示例, 丢包控制的实现方式也可以是对已经进入第一队列的报文进行丢包处理。例如: 从所述至少一个第一队列中的头部开始丢弃报文。又例如: 从所述至少一个第一队列中的尾部开始丢弃报文。需要说明的是, 设备采用对已经进入第一队列的报文进行丢包的拥塞控制方式, 需要该设备支持该丢包方式, 或者, 需要在设备中增加配置信息使得该设备支持该丢包方式。

作为又一个示例, 丢包控制的实现方式也可以对待进入第一队列的报文在入队之前进行丢包处理, 同时, 对已经进入第一队列的报文进行丢包处理。例如: 如图 4 所示, 假设设备的缓存中, S102 中实施拥塞控制的至少一个第一队列包括第一队列 A, 该第一队列 A 包括报文 1、报文 2、...、报文 10, 此时, 可以将即将进入该第一队列 A 的报文 11 和报文 12 进行丢包处理, 而且, 从第一队列的头部 (即从报文 1) 开始进行丢包处理, 从第一队列的尾部 (即从报文 10) 开始进行丢包处理。这样, 可以更加快速的释放缓存中的存储空间, 提高该拥塞控制的效率。

需要说明的是, 如果 S102 中实施拥塞控制的至少一个第一队列包括 idle 队列, 由于 idle 队列自身的特点, 即, idle 队列中无入队和出队操作, 所以, 对该 idle 队列仅可以采用

丢包控制的方式进行拥塞控制，直至将该 idle 队列中的报文全部丢弃。

可见，在本申请实施例中，基于设备缓存的状态设置用于触发拥塞控制的条件，并且，基于缓存中各个队列的实时情况确定进行拥塞控制时待处理的至少一个队列。当确定设备的当前缓存使用信息满足预设的条件时，即可获取该条件对应的待处理的至少一个队列的标识，从而根据所获取的至少一个队列的标识对所述至少一个队列进行拥塞控制。这样，设备不用再基于队列的粒度进行拥塞控制，仅针对总体缓存空间的缓存情况设置用于触发拥塞控制的条件，在触发拥塞控制时设备能够针对性的对该条件对应的队列进行处理，不仅简化配置，提高了缓存空间的利用率，而且在确保不出现缓存溢出问题的同时能够实现高效的拥塞控制，从而提高了设备的网络转发性能。

在本申请实施例中还提供了另一种拥塞控制方法 200，如图 5 所示。实施该方法 200 的设备对缓存中的队列划分了优先级，对于不同的优先级或者相同的优先级中的不同拥塞控制功能，针对性的设置不同的触发拥塞控制的条件，并针对性的记录和实时更新待实施拥塞控制的队列。该方法 200 例如可以包括：

S201，当设备的缓存使用信息满足预设的第一条件时，从缓存中确定该第一条件对应的至少一个第一队列，该第一条件与第一优先级对应，所述至少一个第一队列均属于所述第一优先级。

S202，对所述至少一个第一队列进行拥塞控制。

S203，当缓存使用信息的更新值满足预设的第二条件时，从缓存中确定该第二条件对应的至少一个第二队列，该第二条件与第二优先级对应，所述至少一个第二队列均属于所述第二优先级。

S204，对所述至少一个第二队列进行拥塞控制。

需要说明的是，上述 S201~S202 的实现方式以及效果可以参见方法 100 中 S101~S102 的相关描述。上述 S203~S204 的实现方式以及效果也可以参见方法 100 中 S101~S102 的相关描述。

其中，缓存的使用信息的更新值和 S201 中的缓存的使用信息分别是两个不同时刻用于体现缓存的使用情况的不同使用信息，两者的大小关系在本申请实施例中不作具体限定。例如：缓存的使用信息可以是第一时刻的缓存占用率 95%，缓存的使用信息的更新值可以是第二时刻的缓存占用率 75%；又例如：缓存的使用信息可以是第一时刻的缓存占用率 75%，缓存的使用信息的更新值可以是第二时刻的缓存占用率 95%。

在一些可能的实现方式中，第二优先级可以和第一优先级不同。

如果第二优先级高于第一优先级，作为一个示例，缓存的使用信息为缓存占用率，则，S203 中的缓存的使用信息的更新值大于 S201 中的缓存的使用信息；S201 中的第一条条件可以包括缓存占用率大于第一阈值，S203 中的第二条条件可以包括所述缓存占用率的更新值大于第二阈值，其中，第二阈值大于第一阈值。如图 6a 所示，第一阈值为 70%，第二阈值为 90%，缓存占用率为 75%，缓存占用率的更新值为 95%。作为另一个示例，缓存的使用信息为缓存剩余率，则，S203 中的缓存的使用信息的更新值小于 S201 中的缓存的使用信息；S201 中的第一条条件可以包括缓存剩余率小于第一阈值，S203 中的第二条条件包括缓

存剩余率的更新值小于第二阈值，其中，第二阈值小于第一阈值。如图 6b 所示，第一阈值为 30%，第二阈值为 10%，缓存剩余率为 25%，缓存剩余率的更新值为 5%。

如果第二优先级低于第一优先级，作为一个示例，缓存的使用信息为缓存占用率，则，S203 中的缓存的使用信息的更新值小于 S201 中的缓存的使用信息；S201 中的第一条
5 件可以包括缓存占用率大于第一阈值，S203 中的第二条件可以包括所述缓存占用率的更新值大于第二阈值，其中，第一阈值大于第二阈值。如图 6c 所示，第二阈值为 70%，第一阈值为 90%，缓存占用率为 95%，缓存占用率的更新值为 75%。作为另一个示例，缓存的使用信息为缓存剩余率，则，S203 中的缓存的使用信息的更新值大于 S201 中的缓存的使用信息；S201 中的第一条
10 件可以包括缓存剩余率小于第一阈值，S203 中的第二条件包括缓存剩余率的更新值小于第二阈值，其中，第一阈值小于第二阈值。如图 6d 所示，第二阈值为 30%，第一阈值为 10%，缓存剩余率为 5%，缓存剩余率的更新值为 25%。

举例来说，假设设备按照服务质量（英文：Quality of Service，简称：QoS）将队列划分为 8 个优先级：优先级 0~优先级 7，其中，优先级 7 高于优先级 0，缓存的使用信息为缓存占用率，如图 7a 所示，优先级 0~优先级 7 对应的阈值 1~阈值 8 分别为 20%、30%、
15 40%、50%、60%、70%、80%和 90%，该设备中还记录有待被实施拥塞控制的队列集合 0~队列集合 7，队列集合 0~队列集合 7 分别与优先级 0~优先级 7 对应。当第一时刻缓存的缓存占用率为 25%时，对优先级 0 对应的队列集合 0 中的队列进行拥塞控制；当第二时刻该缓存的缓存占用率为 85%时，拥塞控制可以包括：对优先级 0 对应的队列集合 0 中的队列进行拥塞控制，以及对优先级 6 对应的队列集合 6 中的队列进行拥塞控制。

该实现方式中，通过对缓存中的队列划分不同的优先级，在确定缓存存在溢出风险时，如果缓存中的拥塞程度较轻，则先对优先级较低的队列进行拥塞控制，在拥塞程度较为严重时才对优先级较高的队列进行拥塞控制，这样，将优先级和触发拥塞控制的条件对应的关联设置，先对优先级较低的队列进行拥塞控制的方式，能够确保优先级较高的流量被设备有效的转发，提高了该设备的网络转发性能，提升了用户使用该设备的体验。

在另一些可能的实现方式中，对于第二优先级可以和第一优先级为相同的优先级场景，或者，未划分优先级的场景，S202 和 S204 中的第一条条件和第二条条件可以是基于触发不同拥塞控制方式的时机先后而对应设置的。例如：对于拥塞控制中的反压控制方式，如果同时使能了 ECN 功能和 PFC 功能，而采用 PFC 功能是阻止上游设备发送报文，是无损流量转发等场景中采用的最后拥塞控制手段，如果过早的触发 PFC 功能很可能导致非拥塞
30 状态的出端口无法保持线速，大大降低网络的性能，因此，在确定缓存存在溢出风险时，对于同一个优先级，如果缓存中的拥塞程度较轻，则先触发 ECN 功能对该优先级对应的待处理的 OQ 进行拥塞控制，在拥塞程度较为严重时才触发 PFC 功能对该优先级对应的待处理的 IQ 进行拥塞控制。

对于缓存的使用信息为缓存占用率的情况，一种情况下，如果 S203 中的缓存的使用信息的更新值大于 S201 中的缓存的使用信息，则 S201 中的第一条条件可以包括缓存占用率大于第一阈值，S203 中的第二条件可以包括所述缓存占用率的更新值大于第二阈值，其中，第二阈值大于第一阈值，例如可以参见上述图 6a。另一种情况下，如果 S203 中的缓

存的使用信息的更新值小于 S201 中的缓存使用信息, 则 S201 中的第一条条件可以包括缓存占用率大于第一阈值, S203 中的第二条条件可以包括所述缓存占用率的更新值大于第二阈值, 其中, 第一阈值大于第二阈值, 例如可以参见上述图 6c。

对于缓存的使用信息为缓存剩余率的情况, 一种情况下, 如果 S203 中的缓存的使用信息的更新值小于 S201 中的缓存的使用信息, 则 S201 中的第一条条件可以包括缓存剩余率小于第一阈值, S203 中的第二条条件包括缓存剩余率的更新值小于第二阈值, 其中, 第二阈值小于第一阈值, 例如可以参见上述图 6b。另一种情况下, 如果 S203 中的缓存的使用信息的更新值大于 S201 中的缓存的使用信息, 则 S201 中的第一条条件可以包括缓存剩余率小于第一阈值, S203 中的第二条条件包括缓存剩余率的更新值小于第二阈值, 其中, 第一阈值小于第二阈值, 例如可以参见上述图 6d。

需要说明的是, 该实现方式中, 设备需要记录不同的拥塞控制方式所适用的待被实施拥塞控制的队列, 例如: 对于同时使能 ECN 功能和 PFC 功能的设备, 该设备需要记录 ECN 控制方式对应的待被实施拥塞控制的 OQ, 以及 PFC 控制方式对应的待被实施拥塞控制的 IQ。

举例来说, 假设设备中同时使能 ECN 功能和 PFC 功能, 该设备按照 QoS 将队列划分为 8 个优先级: 优先级 0~优先级 7, 其中, 优先级 7 高于优先级 0, 缓存的使用信息为缓存剩余率, 如图 7b 所示, 优先级 0~优先级 7 对应的阈值 1~阈值 16 分别为 20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%, 其中, 阈值 1、阈值 3、阈值 5、...、阈值 15 分别为优先级 0~优先级 7 对应的 ECN 阈值, 阈值 2、阈值 4、阈值 6、...、阈值 16 分别为优先级 0~优先级 7 对应的 PFC 阈值, 该设备中还记录有待被实施拥塞控制的 IQ 集合 0~IQ 集合 7 以及 OQ 集合 0~OQ 集合 7, IQ 集合 0~IQ 集合 7 分别与优先级 0~优先级 7 对应, OQ 集合 0~OQ 集合 7 分别与优先级 0~优先级 7 对应。当第一时刻缓存的缓存占用率为 22%时, 对优先级 0 对应的 OQ 集合 0 中的 OQ 进行 ECN 控制, 即, 将 OQ 集合 0 中 OQ 的出队报文添加 ECN 标记; 当第二时刻该缓存的缓存占用率为 39%时, 不仅将 OQ 集合 0 中 OQ 的出队报文添加 ECN 标记, 而且对优先级 1 对应的 IQ 集合 1 中的 IQ 进行 PFC 控制, 其中, 对优先级 1 对应的 IQ 集合 1 中的 IQ 进行 PFC 控制的方式例如可以包括: 通过 IQ 集合 1 中 IQ 对应的入端口向上游设备 (或上游缓存) 发送 PFC 帧。这样, 将 ECN 控制方式和 PFC 控制方式分别与不同的触发条件关联设置, 能够确保未划分优先级或同一优先级的场景中, 先触发 ECN 控制后触发 PFC 控制, 提高了拥塞控制的合理性。

该实现方式中, 在同时使能多种拥塞控制方式且确定缓存存在溢出风险时, 合理设置多种拥塞控制方式触发的先后顺序, 并按照拥塞控制方式被触发的先后顺序设置对应的触发条件, 这样, 如果缓存中的拥塞程度较轻, 则利用所确定的需要先触发的拥塞控制方式进行拥塞控制, 在拥塞程度较为严重时才利用所确定的需要后触发的拥塞控制方式进行拥塞控制, 将拥塞控制方式和触发拥塞控制的条件对应的关联设置, 能够确保队列的拥塞控制更加合理, 提高了该设备的网络转发性能, 提升了用户使用该设备的体验。

需要说明的是, 对于用于触发拥塞控制的多个不同条件触发的拥塞控制, 可以认为是

解耦的过程，即，当设备的缓存的使用信息满足预设的第一条件时，一直会对第一条件对应的至少一个第一队列进行拥塞控制，直到设备的缓存的使用信息不再满足该第一条件；同理，当设备的缓存的使用信息满足预设的第二条件时，一直会对第二条件对应的至少一个第二队列进行拥塞控制，直到设备的缓存的使用信息不再满足该第二条件，两个拥塞控制的过程不相互干扰。

可见，在本申请实施例中，基于设备缓存的状态预设多个用于触发拥塞控制的条件，并且，基于缓存中各个队列的实时情况确定进行不同拥塞控制时待处理的队列。在不同的时刻，当确定设备的缓存使用信息满足不同的条件时，对该条件对应的待处理的队列进行相应的拥塞控制。这样，设备不用再基于队列的粒度进行拥塞控制，仅针对总体缓存空间的缓存情况设置用于触发不同拥塞控制的不同条件，能够针对性的对不同条件对应的队列进行处理，不仅简化配置，提高了缓存空间的利用率，而且在确保不出现缓存溢出问题的同时能够实现高效、精准的拥塞控制，从而提高了设备的网络转发性能。

图 8 为本申请实施例提供的一种拥塞控制装置的结构示意图。该装置 800 可以是设备，该装置 800 可以包括：第一确定单元 801 和第一控制单元 802。

其中，第一确定单元 801 用于当设备缓存的使用信息满足预设的第一条件时，从缓存中确定大于等于队列时延阈值或者大于等于队列长度阈值的至少一个第一队列，第一条件为缓存的使用信息对应的第一阈值。

第一控制单元 802 用于对至少一个第一队列进行拥塞控制。

需要说明的是，该装置 800 用于执行上述图 2 或图 5 所示实施例对应的各个步骤，以按照本申请实施例提供的拥塞控制方法实现拥塞控制。其中，第一确定单元 801 对应上述 S101 或 S201，第一控制单元 802 对应上述 S102 或 S202。

在一些可能的实现方式中，该第一确定单元 801 具体用于确定缓存中多个队列按照队列时延从长到短排序后的前 M 个队列，作为至少一个第一队列，队列时延阈值与第 M 个队列的队列时延相等，M 为正整数；或者，第一确定单元 801，具体用于：确定缓存中多个队列按照队列长度从长到短排序后的前 N 个队列，作为至少一个第一队列，队列长度阈值与第 N 个队列的队列长度相等，N 为正整数。

作为一个示例，第一条件与第一优先级对应，至少一个第一队列属于第一优先级。

该示例下，该第一确定单元 801 具体用于确定第一优先级对应的多个队列按照队列时延从长到短排序后的前 K 个队列，作为至少一个第一队列，队列时延阈值与第 K 个队列的队列时延相等，K 为正整数；或者，第一确定单元 801，具体用于：确定第一优先级对应的多个队列按照队列长度从长到短排序后的前 L 个队列，作为至少一个第一队列，队列长度阈值与第 L 个队列的队列长度相等，L 为正整数。

可以理解的是，缓存的使用信息为缓存占用率，第一条件包括缓存占用率大于第一阈值；或者，缓存使用信息为缓存剩余率，第一条件包括缓存剩余率小于第一阈值。

在另一些可能的实现方式中，该装置 800 还可以包括：第二确定单元和第二控制单元。其中，第二确定单元，用于当缓存的使用信息的更新值满足预设的第二条件时，从缓存中确定第二条件对应的至少一个第二队列，第二条件与第二优先级对应，至少一个第二

队列属于第二优先级，第二条件为缓存的使用信息的更新值对应的第二阈值；第二控制单元，用于对至少一个第二队列进行拥塞控制。

需要说明的是，该装置 800 用于执行上述图 5 所示实施例对应的各个步骤，以按照本申请实施例提供的拥塞控制方法实现拥塞控制。其中，第一确定单元 801 对应上述 S201，
5 第一控制单元 802 对应上述 S202，第二确定单元对应上述 S203，第二控制单元对应上述 S204。

作为一个示例，第二优先级和第一队列所属的第一优先级不同。

作为另一个示例，第二优先级和第一队列所属的第一优先级相同。

在一些可能的实现方式中，缓存的使用信息的更新值为缓存占用率的更新值，第二条件包括缓存占用率的更新值大于第二阈值，第二阈值大于第一阈值；或者，缓存的使用信息的更新值为缓存剩余率的更新值，第二条件包括缓存剩余率的更新值小于第二阈值，第二阈值小于第一阈值。
10

在另一些可能的实现方式中，第一控制单元具体用于：为至少一个第一队列中的出队报文添加显示拥塞通告 ECN 标记，ECN 标记用于指示源端设备降低与第一队列对应的队列中的报文的发送速率。该实现方式中，第二控制单元，具体用于：通过至少一个第二队列的入端口向设备的上游设备发送第一反压信号，第一反压信号用于指示上游设备停止发送报文。
15

可以理解的是，第一控制单元具体用于：对至少一个第一队列进行反压控制或丢包控制。

在一些可能的实现方式中，至少一个第一队列还包括预设时长内无出队报文且无入队报文的队列。
20

需要说明的是，本申请实施例提供的拥塞控制装置 800，对应于上述拥塞控制方法，故，该装置 800 的各种可能的实现方式以及达到的技术效果，可以参照前述图 2 和图 5 对应实施例的相关描述。

此外，本申请实施例还提供了一种设备，如图 9 所示，该设备 900 包括存储器 901 和处理器 902。其中，存储器 901，用于存储程序代码；处理器 902，用于运行程序代码中的指令，使得设备 900 执行前述图 2 和图 5 对应实施例中任意一种实现方式的拥塞控制方法。
25

示例的，该设备 900 还可以包括通信接口，用于接收报文，将报文存入缓存对应的队列中。
30

示例的，该设备 900 的处理器 902，用于运行存储器 901 上保存的程序代码中的指令，使得该设备 900 执行下述拥塞控制方法：

当设备缓存的使用信息满足预设的第一条件时，从所述缓存中确定大于等于队列时延阈值或者大于等于队列长度阈值的至少一个第一队列，所述第一条件为所述缓存的使用信息对应的第一阈值；
35

对所述至少一个第一队列进行拥塞控制。

其中，存储器 901 可以是易失性存储器（volatile memory），例如随机存取存储器

(random-access memory, RAM) ; 或者非易失性存储器 (non-volatile memory) , 例如只读存储器 (read-only memory, ROM) , 快闪存储器 (flash memory) , 硬盘 (hard disk drive, HDD) 或固态硬盘 (solid-state drive, SSD) ; 或者上述种类的存储器的组合, 用于存储可实现本申请方法的程序代码、TSN 域内网络设备的配置文件或其他内容。

5 处理器 902 是控制器的控制中心, 可以是一个中央处理器 (central processing unit, CPU) , 也可以是特定集成电路 (application specific integrated circuit, ASIC) , 或者是被配置成实施本申请实施例的一个或多个集成电路, 例如: 一个或多个数字信号处理器 (digital signal processor, DSP) , 或, 一个或者多个现场可编程门阵列 (field programmable gate array, FPGA) 。

10 通信接口用于和其他设备进行数据交互, 例如: 通过该通信接口从其他设备接收报文, 又例如: 通过该通信接口向其他设备发送报文通信接口可以为以太网接口 (Ethernet) 接口、快速以太 (Fast Ethernet, FE) 接口或千兆以太 (Gigabit Ethernet, GE) 接口。

此外, 本申请实施例还提供了一种计算机程序产品, 当其在计算机上运行时, 使得计算机执行前述图 2 和图 5 对应实施例中任意一种实现方式所述的拥塞控制方法。

15 示例的, 该计算机程序产品, 当其在计算机上运行时, 例如在设备 900 上运行时, 使得计算机执行下述拥塞控制方法:

当设备缓存的使用信息满足预设的第一条件时, 从所述缓存中确定大于等于队列时延阈值或者大于等于队列长度阈值的至少一个第一队列, 所述第一条件为所述缓存的使用信息对应的第一阈值;

20 对所述至少一个第一队列进行拥塞控制。

此外, 本申请实施例还提供了一种计算机可读存储介质, 该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, 使得该计算机执行前述图 2 和图 5 对应实施例中任意一种实现方式所述的拥塞控制方法。

25 示例的, 该计算机可读存储介质, 该计算机可读存储介质中存储有指令或程序代码, 当其在计算机上运行时, 例如在设备 900 上运行时, 使得该计算机执行下述拥塞控制方法:

当设备缓存的使用信息满足预设的第一条件时, 从所述缓存中确定大于等于队列时延阈值或者大于等于队列长度阈值的至少一个第一队列, 所述第一条件为所述缓存的使用信息对应的第一阈值;

30 对所述至少一个第一队列进行拥塞控制。

35 需说明的是, 以上描述的任意装置实施例都仅仅是示意性的, 其中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模块来实现本实施例方案的目的。另外, 本申请提供的设备实施例附图中, 模块之间的连接关系表示它们之间具有通信连接, 具体可以实现为一条或多条通信总线或信号线。本领域普通技术人员在不付出创造性劳动的情况下, 即可以理解并实施。

本申请实施例公开内容所描述的方法或者算法的步骤可以硬件的方式来实现，也可以是由处理器执行软件指令的方式来实现。软件指令可以由相应的软件模块组成，软件模块可以被存放于随机存取存储器（random access memory, RAM）、闪存、只读存储器（read only memory, ROM）、可擦除可编程只读存储器（erasable programmable ROM, EPROM）、电可擦可编程只读存储器（electrically EPROM, EEPROM）、硬盘、移动硬盘、光盘或者本领域熟知的任何其它形式的存储介质中。一种示例性的存储介质耦合至处理器，从而使处理器能够从该存储介质读取信息，且可向该存储介质写入信息。当然，存储介质也可以是处理器的组成部分。处理器和存储介质可以位于ASIC中。另外，该ASIC可以位于核心网接口设备中。当然，处理器和存储介质也可以作为分立组件存在于核心网接口设备中。

本说明书中的各个实施例均采用递进的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于装置实施例而言，由于其基本相似于方法实施例，所以描述得比较简单，相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的，其中作为分离部件说明的模块可以是或者也可以不是物理上分开的，作为模块显示的部件可以是或者也可以不是物理模块，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模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下，即可以理解并实施。

以上所述仅是本申请示例性的实施方式，并非用于限定本申请的保护范围。

权 利 要 求

1.一种拥塞控制方法，其特征在于，包括：

当设备缓存的使用信息满足预设的第一条件时，从所述缓存中确定大于等于队列时延
5 阈值或者大于等于队列长度阈值的至少一个第一队列，所述第一条件包括所述缓存的使用
信息对应的第一阈值；

对所述至少一个第一队列进行拥塞控制。

2.根据权利要求1所述的方法，其特征在于，所述从所述缓存中确定大于等于队列时延
10 阈值的至少一个第一队列包括：

确定所述缓存中多个队列按照队列时延从长到短排序后的前M个队列，作为所述至少
10 一个第一队列，所述队列时延阈值与第M个队列的队列时延相等，所述M为正整数；

所述从所述缓存中确定大于等于队列长度阈值的至少一个第一队列包括：

确定所述缓存中多个队列按照队列长度从长到短排序后的前N个队列，作为所述至少
一个第一队列，所述队列长度阈值与第N个队列的队列长度相等，所述N为正整数。

3.根据权利要求1所述的方法，其特征在于，所述第一条件与第一优先级对应，所述至
15 少一个第一队列属于所述第一优先级。

4.根据权利要求3所述的方法，其特征在于，所述从所述缓存中确定大于队列时延阈值
的至少一个第一队列包括：

确定所述第一优先级对应的多个队列按照队列时延从长到短排序后的前K个队列，作
20 为所述至少一个第一队列，所述队列时延阈值与第K个队列的队列时延相等，所述K为正
整数；

所述从所述缓存中确定大于队列长度阈值的至少一个第一队列包括：

确定所述第一优先级对应的多个队列按照队列长度从长到短排序后的前L个队列，作
为所述至少一个第一队列，所述队列长度阈值与第L个队列的队列长度相等，所述L为正
整数。

5.根据权利要求1-4任一项所述的方法，其特征在于，

所述缓存的使用信息为缓存占用率，所述第一条件包括所述缓存占用率大于所述第一
25 阈值；或者，

所述缓存使用信息为缓存剩余率，所述第一条件包括所述缓存剩余率小于所述第一阈
值。

6.根据权利要求1-5任一项所述的方法，其特征在于，所述方法还包括：

当所述缓存的使用信息的更新值满足预设的第二条件时，从所述缓存中确定所述第二
30 条件对应的至少一个第二队列，所述第二条件与第二优先级对应，所述至少一个第二队列
属于所述第二优先级，所述第二条件为所述缓存的使用信息的更新值对应的第二阈值；

对所述至少一个第二队列进行拥塞控制。

7.根据权利要求6所述的方法，其特征在于，所述第二优先级和所述第一队列所属的第
35 一优先级不同。

8.根据权利要求6所述的方法，其特征在于，所述第二优先级和所述第一队列所属的第

一优先级相同。

9.根据权利要求 6-8 任一项所述的方法,其特征在于,

所述缓存的使用信息的更新值为缓存占用率的更新值,所述第二条件包括所述缓存占用率的更新值大于所述第二阈值,所述第二阈值大于所述第一阈值;或者,

5 所述缓存的使用信息的更新值为缓存剩余率的更新值,所述第二条件包括所述缓存剩余率的更新值小于所述第二阈值,所述第二阈值小于所述第一阈值。

10.根据权利要求 9 所述的方法,其特征在于,所述对所述至少一个第一队列进行拥塞控制,包括:

10 为所述至少一个第一队列中的出队报文添加显示拥塞通告 ECN 标记,所述 ECN 标记用于指示源端设备降低与所述第一队列对应的队列中的报文的发送速率。

11.根据权利要求 10 所述的方法,其特征在于,所述对所述至少一个第二队列进行拥塞控制,包括:

通过所述至少一个第二队列的入端口向所述设备的上游设备发送反压信号,所述反压信号用于指示所述上游设备停止发送报文。

15 12.根据权利要求 1-11 任一项所述的方法,其特征在于,所述对所述至少一个第一队列进行拥塞控制,包括:

对所述至少一个第一队列进行反压控制或丢包控制。

13.根据权利要求 1-12 任一项所述的方法,其特征在于,所述至少一个第一队列还包括预设时长内无出队报文且无入队报文的队列。

20 14.一种拥塞控制装置,其特征在于,包括:

第一确定单元,用于当设备缓存的使用信息满足预设的第一条件时,从所述缓存中确定大于等于队列时延阈值或者大于等于队列长度阈值的至少一个第一队列,所述第一条件包括所述缓存的使用信息对应的第一阈值;

第一控制单元,用于对所述至少一个第一队列进行拥塞控制。

25 15.根据权利要求 14 所述的装置,其特征在于,所述第一确定单元,具体用于:

确定所述缓存中多个队列按照队列时延从长到短排序后的前 M 个队列,作为所述至少一个第一队列,所述队列时延阈值与第 M 个队列的队列时延相等,所述 M 为正整数;或者,

30 确定所述缓存中多个队列按照队列长度从长到短排序后的前 N 个队列,作为所述至少一个第一队列,所述队列长度阈值与第 N 个队列的队列长度相等,所述 N 为正整数。

16.根据权利要求 14 所述的装置,其特征在于,所述第一条件与第一优先级对应,所述至少一个第一队列属于所述第一优先级。

17.根据权利要求 16 所述的装置,其特征在于,所述第一确定单元,具体用于:

35 确定所述第一优先级对应的多个队列按照队列时延从长到短排序后的前 K 个队列,作为所述至少一个第一队列,所述队列时延阈值与第 K 个队列的队列时延相等,所述 K 为正整数;或者,

确定所述第一优先级对应的多个队列按照队列长度从长到短排序后的前 L 个队列,作

为所述至少一个第一队列，所述队列长度阈值与第 L 个队列的队列长度相等，所述 L 为正整数。

18.根据权利要求 14-17 任一项所述的装置，其特征在于，

5 所述缓存的使用信息为缓存占用率，所述第一条件包括所述缓存占用率大于所述第一阈值；或者，

所述缓存使用信息为缓存剩余率，所述第一条件包括所述缓存剩余率小于所述第一阈值。

19.根据权利要求 14-18 任一项所述的装置，其特征在于，所述装置还包括：

10 第二确定单元，用于当所述缓存的使用信息的更新值满足预设的第二条件时，从所述缓存中确定所述第二条件对应的至少一个第二队列，所述第二条件与第二优先级对应，所述至少一个第二队列属于所述第二优先级，所述第二条件为所述缓存的使用信息的更新值对应的第二阈值；

第二控制单元，用于对所述至少一个第二队列进行拥塞控制。

15 20.根据权利要求 19 所述的装置，其特征在于，所述第二优先级和所述第一队列所属的第一优先级不同。

21.根据权利要求 19 所述的装置，其特征在于，所述第二优先级和所述第一队列所属的第一优先级相同。

22.根据权利要求 19-21 任一项所述的装置，其特征在于，

20 所述缓存的使用信息的更新值为缓存占用率的更新值，所述第二条件包括所述缓存占用率的更新值大于所述第二阈值，所述第二阈值大于所述第一阈值；或者，

所述缓存的使用信息的更新值为缓存剩余率的更新值，所述第二条件包括所述缓存剩余率的更新值小于所述第二阈值，所述第二阈值小于所述第一阈值。

23.根据权利要求 22 所述的装置，其特征在于，所述第一控制单元，具体用于：

25 为所述至少一个第一队列中的出队报文添加显示拥塞通告 ECN 标记，所述 ECN 标记用于指示源端设备降低与所述第一队列对应的队列中的报文的发送速率。

24.根据权利要求 23 所述的装置，其特征在于，所述第二控制单元，具体用于：

通过所述至少一个第二队列的入端口向所述设备的上游设备发送反压信号，所述反压信号用于指示所述上游设备停止发送报文。

30 25.根据权利要求 14-24 任一项所述的装置，其特征在于，所述第一控制单元，具体用于：

对所述至少一个第一队列进行反压控制或丢包控制。

26.根据权利要求 14-25 任一项所述的装置，其特征在于，所述至少一个第一队列还包括预设时长内无出队报文且无入队报文的队列。

27.一种设备，其特征在于，所述设备包括存储器和处理器；

35 所述存储器，用于存储程序代码；

所述处理器，用于运行所述程序代码中的指令，使得所述设备执行以上权利要求 1-13 任一项所述的拥塞控制方法。

28.一种计算机可读存储介质，其特征在于，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，使得所述计算机执行以上权利要求 1-13 任一项所述的拥塞控制方法。

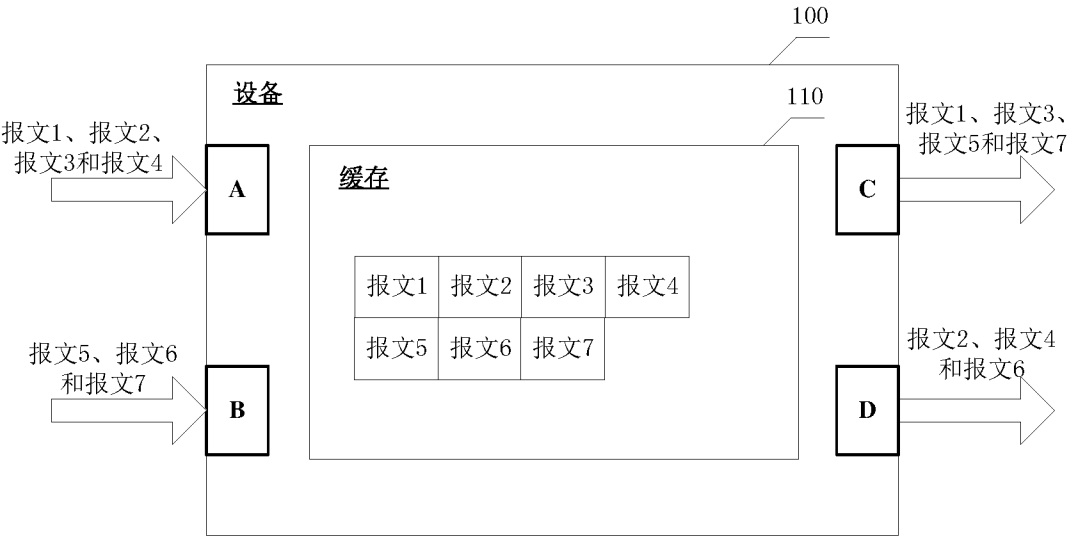


图 1a

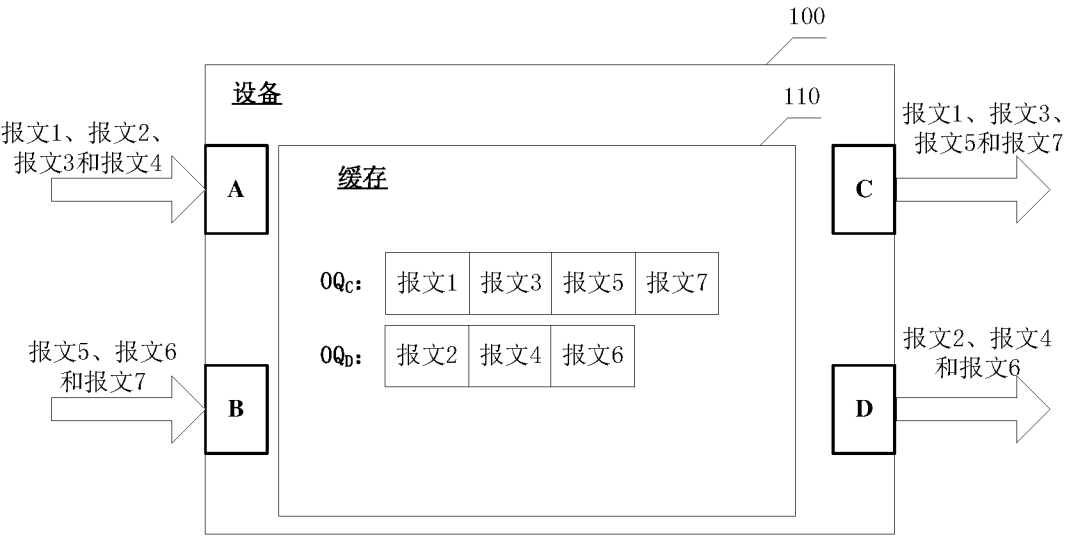


图 1b

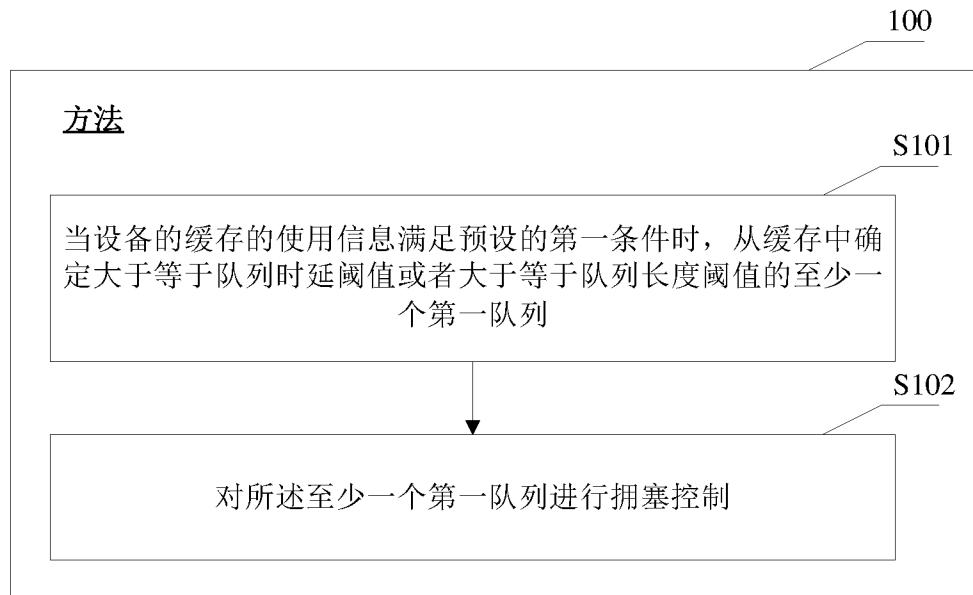


图 2

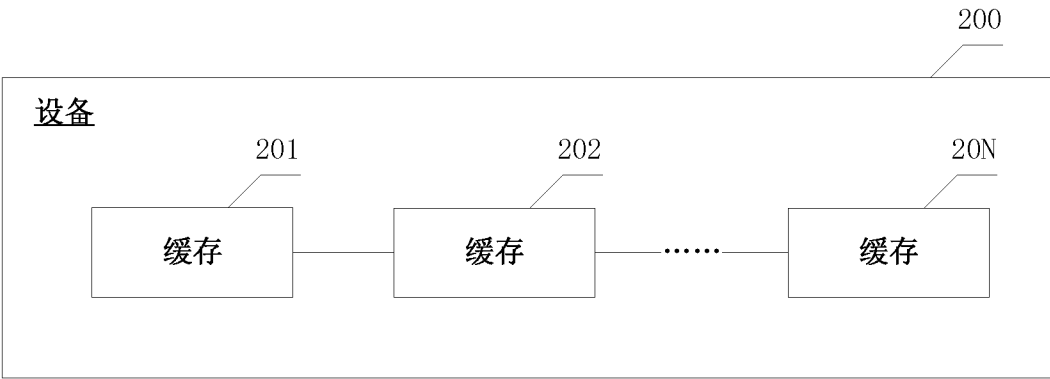


图 3

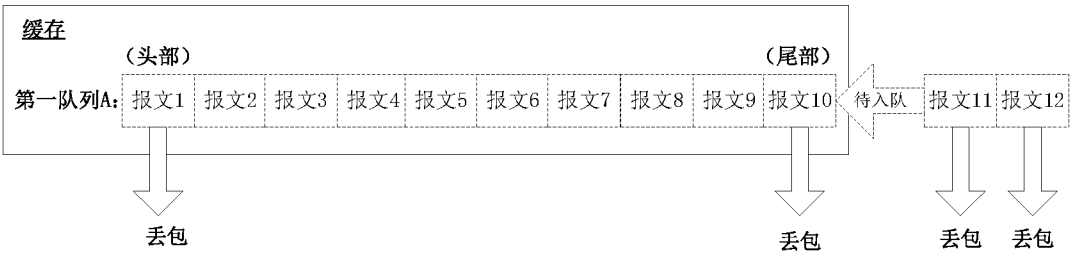


图 4

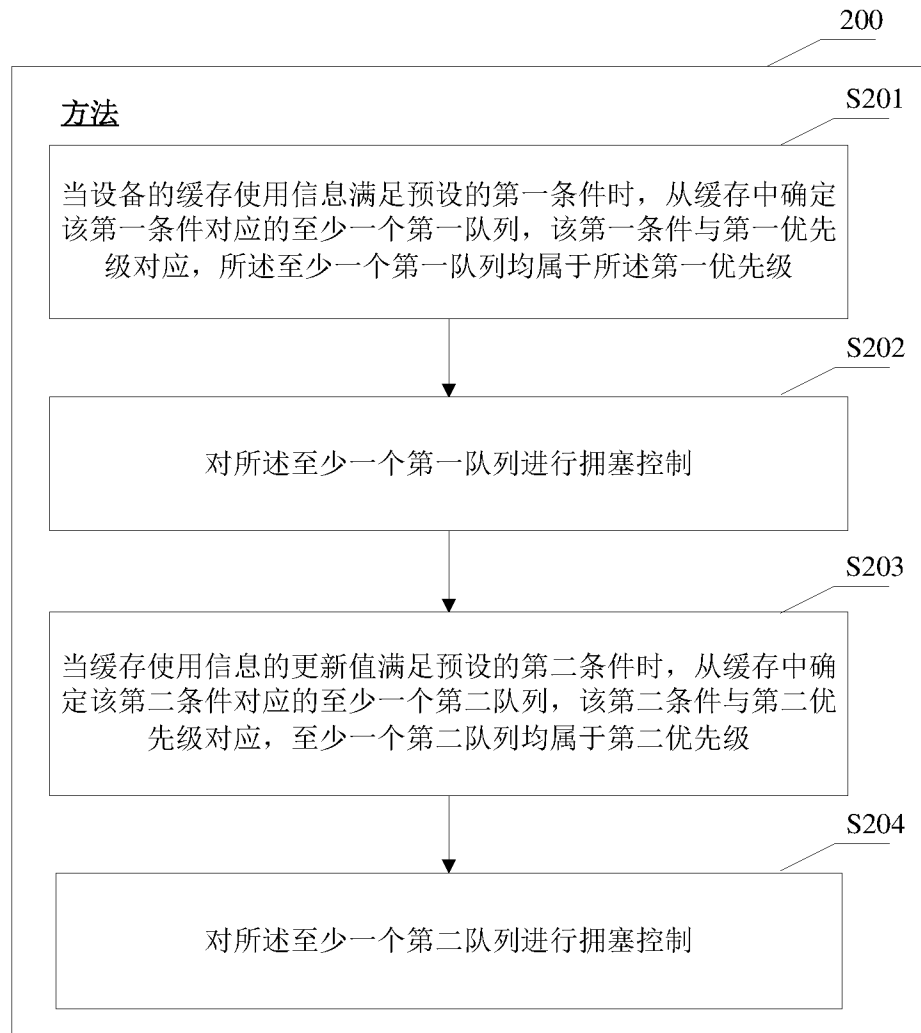


图 5

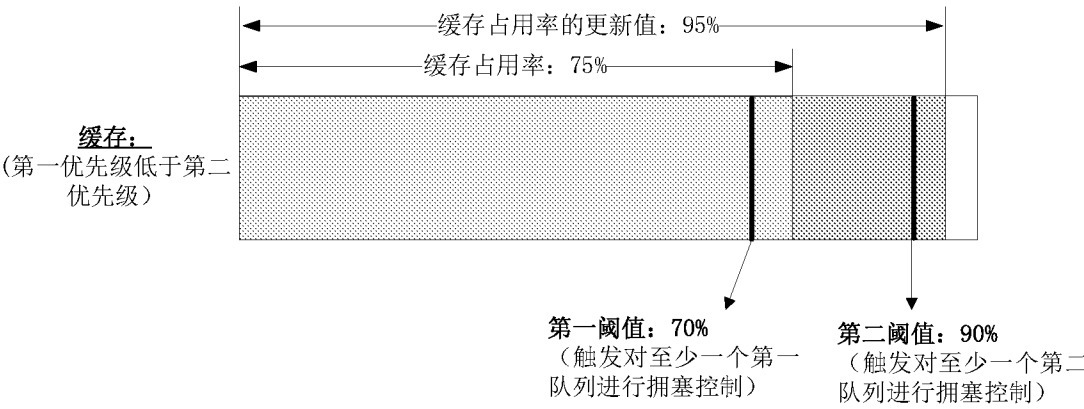


图 6a

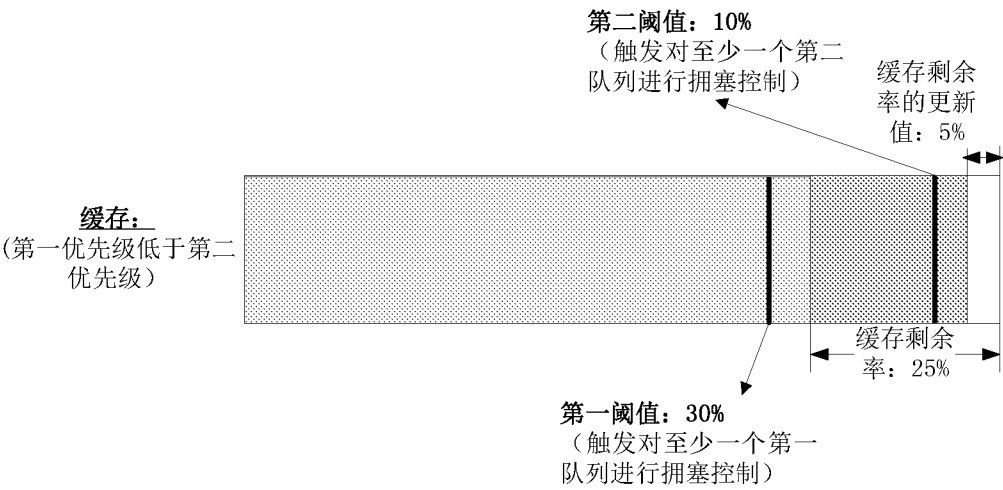


图 6b

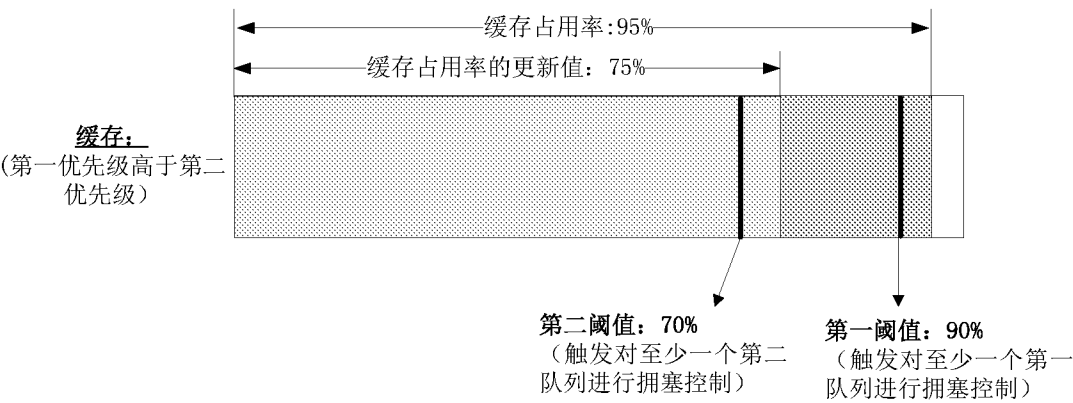


图 6c

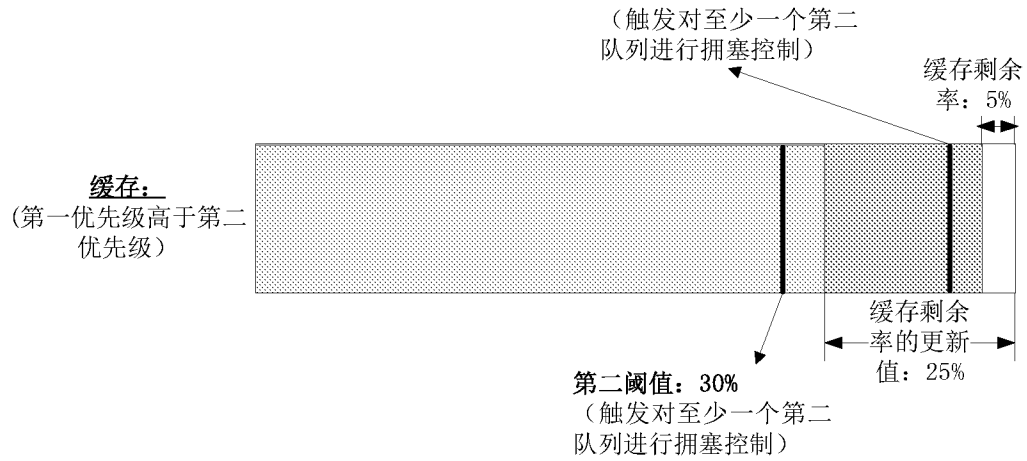


图 6d

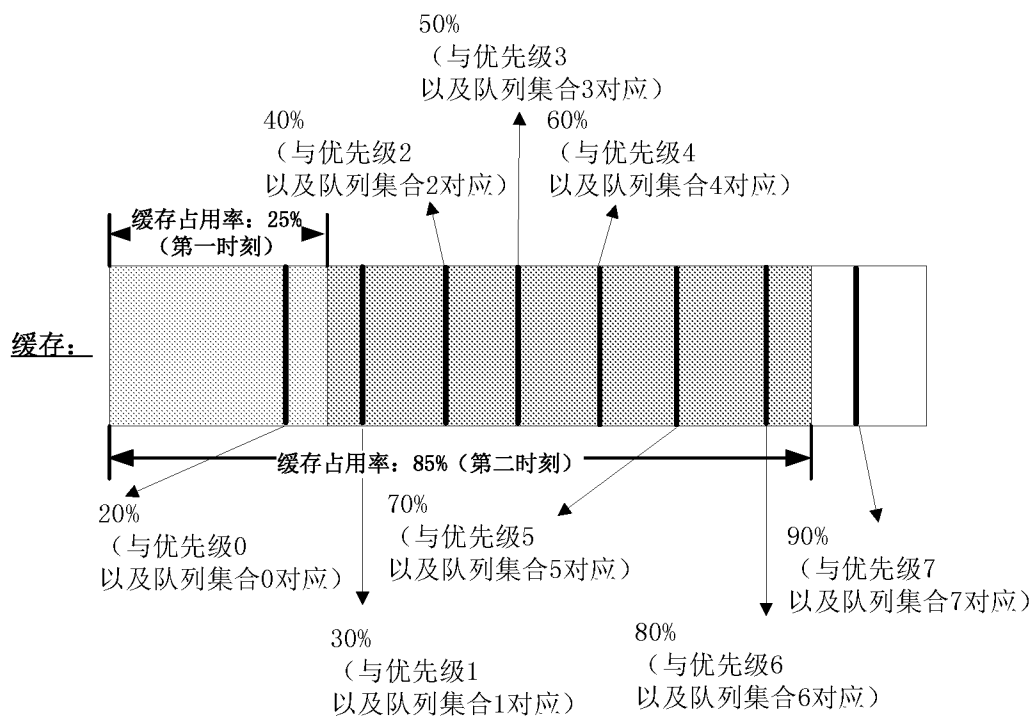


图 7a

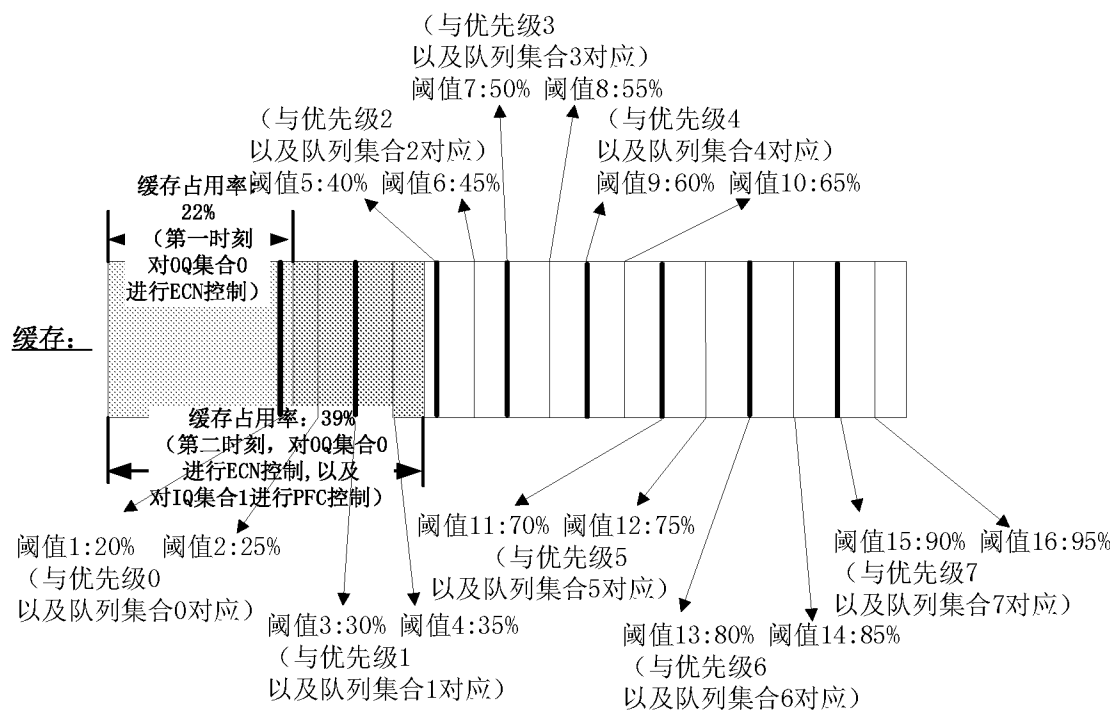


图 7b

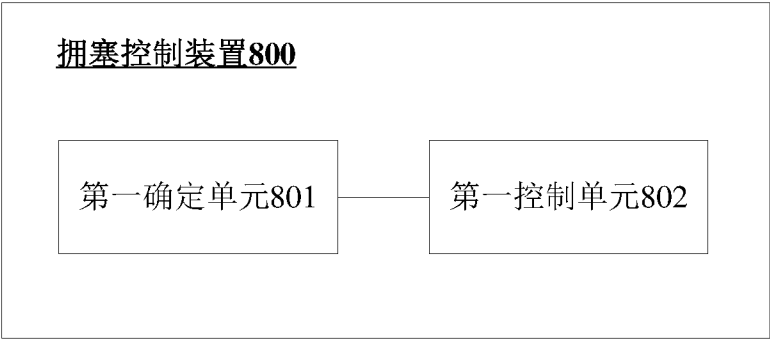


图 8

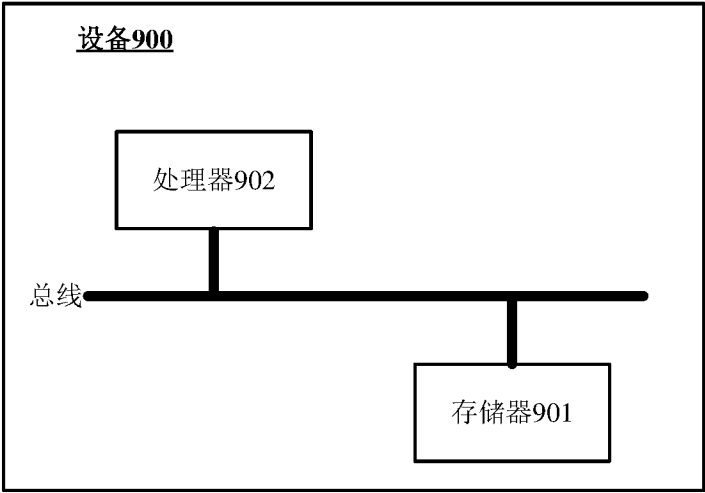


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/081462

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/801(2013.01)i; H04L 12/803(2013.01)i; H04L 12/835(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)

CNPAT, CNKI, WPI, EPODOC: 缓存, 拥塞, 控制, 占用率, 剩余率, 队列, 时延, 长度, 丢包, 丢弃, 反压, 降低, 速率, 停止, 发送, buffer, cache, congestion, control, occupancy, remain, queue, delay, length, packet, drop, reduce, decrease, lower, rate, stop

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 109660468 A (SANECHIPS TECHNOLOGY CO., LTD.) 19 April 2019 (2019-04-19) description, paragraphs [0029]-[0094], and figures 1-4	1-28
A	US 2017272372 A1 (MELLANOX TECHNOLOGIES TLV LTD.) 21 September 2017 (2017-09-21) entire document	1-28
A	CN 108206787 A (BEIJING HUAWEI DIGITAL TECHNOLOGY CO., LTD.) 26 June 2018 (2018-06-26) entire document	1-28
A	CN 110011924 A (SANECHIPS TECHNOLOGY CO., LTD.) 12 July 2019 (2019-07-12) entire document	1-28
A	CN 110278157 A (HUAWEI TECHNOLOGIES CO., LTD.) 24 September 2019 (2019-09-24) entire document	1-28



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

01 June 2021

Date of mailing of the international search report

17 June 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
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/CN2021/081462

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109660468	A	19 April 2019	None			
US	2017272372	A1	21 September 2017	CN	107204931	A	26 September 2017
CN	108206787	A	26 June 2018	None			
CN	110011924	A	12 July 2019	None			
CN	110278157	A	24 September 2019	EP	3758315	A1	30 December 2020
				US	2020412660	A1	31 December 2020
				WO	2019174536	A1	19 September 2019

国际检索报告

国际申请号

PCT/CN2021/081462

A. 主题的分类 H04L 12/801(2013.01)i; H04L 12/803(2013.01)i; H04L 12/835(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)) CNPAT, CNKI, WPI, EPDOC: 缓存, 拥塞, 控制, 占用率, 剩余率, 队列, 时延, 长度, 丢包, 丢弃, 反压, 降低, 速率, 停止, 发送, buffer, cache, congestion, control, occupancy, remain, queue, delay, length, packet, drop, reduce, decrease, lower, rate, stop																				
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>X</td> <td>CN 109660468 A (深圳市中兴微电子有限公司) 2019年 4月 19日 (2019 - 04 - 19) 说明书第[0029]-[0094]段, 图1-4</td> <td>1-28</td> </tr> <tr> <td>A</td> <td>US 2017272372 A1 (MELLANOX TECHNOLOGIES TLV LTD.) 2017年 9月 21日 (2017 - 09 - 21) 全文</td> <td>1-28</td> </tr> <tr> <td>A</td> <td>CN 108206787 A (北京华为数字技术有限公司) 2018年 6月 26日 (2018 - 06 - 26) 全文</td> <td>1-28</td> </tr> <tr> <td>A</td> <td>CN 110011924 A (深圳市中兴微电子有限公司) 2019年 7月 12日 (2019 - 07 - 12) 全文</td> <td>1-28</td> </tr> <tr> <td>A</td> <td>CN 110278157 A (华为技术有限公司) 2019年 9月 24日 (2019 - 09 - 24) 全文</td> <td>1-28</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	X	CN 109660468 A (深圳市中兴微电子有限公司) 2019年 4月 19日 (2019 - 04 - 19) 说明书第[0029]-[0094]段, 图1-4	1-28	A	US 2017272372 A1 (MELLANOX TECHNOLOGIES TLV LTD.) 2017年 9月 21日 (2017 - 09 - 21) 全文	1-28	A	CN 108206787 A (北京华为数字技术有限公司) 2018年 6月 26日 (2018 - 06 - 26) 全文	1-28	A	CN 110011924 A (深圳市中兴微电子有限公司) 2019年 7月 12日 (2019 - 07 - 12) 全文	1-28	A	CN 110278157 A (华为技术有限公司) 2019年 9月 24日 (2019 - 09 - 24) 全文	1-28
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
X	CN 109660468 A (深圳市中兴微电子有限公司) 2019年 4月 19日 (2019 - 04 - 19) 说明书第[0029]-[0094]段, 图1-4	1-28																		
A	US 2017272372 A1 (MELLANOX TECHNOLOGIES TLV LTD.) 2017年 9月 21日 (2017 - 09 - 21) 全文	1-28																		
A	CN 108206787 A (北京华为数字技术有限公司) 2018年 6月 26日 (2018 - 06 - 26) 全文	1-28																		
A	CN 110011924 A (深圳市中兴微电子有限公司) 2019年 7月 12日 (2019 - 07 - 12) 全文	1-28																		
A	CN 110278157 A (华为技术有限公司) 2019年 9月 24日 (2019 - 09 - 24) 全文	1-28																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
<table border="0"> <tr> <td> * 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 </td> <td> “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 </td> </tr> </table>			* 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																
* 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 2021年 6月 1日		国际检索报告邮寄日期 2021年 6月 17日																		
ISA/CN的名称和邮寄地址 中国国家知识产权局(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 王健 电话号码 86-(10)-53961733																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/081462

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	109660468	A	2019年 4月 19日	无	
US	2017272372	A1	2017年 9月 21日	CN	107204931 A 2017年 9月 26日
CN	108206787	A	2018年 6月 26日	无	
CN	110011924	A	2019年 7月 12日	无	
CN	110278157	A	2019年 9月 24日	EP	3758315 A1 2020年 12月 30日
				US	2020412660 A1 2020年 12月 31日
				WO	2019174536 A1 2019年 9月 19日