



(12) 发明专利申请

(10) 申请公布号 CN 104219150 A

(43) 申请公布日 2014. 12. 17

(21) 申请号 201410447480. 6

(22) 申请日 2014. 09. 03

(71) 申请人 杭州华三通信技术有限公司
地址 310052 浙江省杭州市滨江区长河路
466 号

(72) 发明人 程锦平

(74) 专利代理机构 北京博思佳知识产权代理有限公司 11415
代理人 林祥

(51) Int. Cl.
H04L 12/741 (2013. 01)

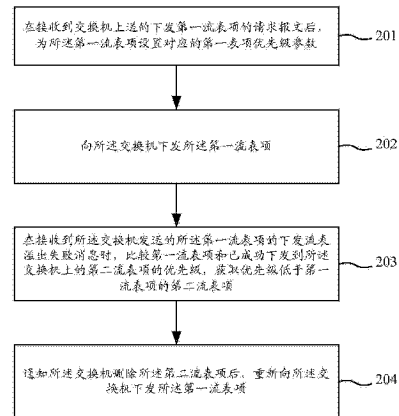
权利要求书2页 说明书7页 附图3页

(54) 发明名称

流表下发方法及装置

(57) 摘要

本发明提供一种流表下发方法及装置,应用于软件定义网络 SDN 控制器,该方法包括:在接收到交换机上送的下发第一流表项的请求报文后,为所述第一流表项设置对应的第一表项优先级参数;向所述交换机下发所述第一流表项;在接收到所述交换机发送的所述第一流表项的下发表项溢出失败消息时,通过比较第一流表项和已成功下发到所述交换机上的第二流表项的优先级,获取优先级低于第一流表项的第二流表项;通知所述交换机删除所述第二流表项后,重新向所述交换机下发所述第一流表项。在交换机硬件资源用尽的情况下,本发明通过设置业务优先级,删除低优先级业务,保证高优先级业务的正常转发,提高 SDN 网络的业务传输质量。



1. 一种流表下发方法,应用于软件定义网络 SDN 控制器,其特征在于,该方法包括:
 - 在接收到交换机上送的下发第一流表项的请求报文后,为所述第一流表项设置对应的第一表项优先级参数;
 - 向所述交换机下发所述第一流表项;
 - 在接收到所述交换机发送的所述第一流表项的下发流表溢出失败消息时,通过比较第一流表项和已成功下发到所述交换机上的第二流表项的优先级,获取优先级低于第一流表项的第二流表项;
 - 通知所述交换机删除所述第二流表项后,重新向所述交换机下发所述第一流表项。
2. 如权利要求 1 所述的方法,其特征在于,所述为所述第一流表项设置对应的第一表项优先级参数,包括:
 - 预先构建业务优先级列表;
 - 根据所述第一流表项的请求报文中携带的业务标识信息,查找所述业务优先级列表;
 - 根据查找结果,获取所述请求报文对应的业务优先级参数;
 - 根据所述业务优先级参数,设置对应的所述第一表项优先级参数。
3. 如权利要求 2 所述的方法,其特征在于,所述根据所述第一流表项的请求报文中携带的业务标识信息,查找预先设置的业务优先级列表,包括:
 - 根据所述请求报文携带的 IP 地址和 MAC 地址,查询预设的第一业务优先级列表;
 - 若所述第一业务优先级列表中无匹配表项,则根据所述请求报文携带的优先级字段信息,查询预设的第二业务优先级列表。
4. 如权利要求 1 所述的方法,其特征在于,所述通过比较第一流表项和已成功下发到所述交换机上的第二流表项的优先级,获取优先级低于第一流表项的第二流表项,包括:
 - 将所述第一表项优先级参数与其保存的已成功下发到所述交换机上的流表项的表项优先级参数进行比较;
 - 根据比较结果,从表项优先级低于所述第一表项优先级的流表项中,选择优先级最低的流表项作为所述第二流表项。
5. 如权利要求 4 所述的方法,其特征在于,所述选择优先级最低的流表项作为所述第二流表项,包括:
 - 若存在多个所述优先级最低的流表项,则从所述多个优先级最低的流表项中选择报文匹配最少的流表项作为所述第二流表项。
6. 一种流表下发装置,应用于软件定义网络 SDN 控制器,其特征在于,该装置包括:
 - 表项优先级设置单元,用于在接收到交换机上送的下发第一流表项的请求报文后,为所述第一流表项设置对应的第一表项优先级参数;
 - 流表项下发单元,用于向所述交换机下发所述第一流表项;
 - 流表项获取单元,用于在接收到所述交换机发送的所述第一流表项的下发流表溢出失败消息时,比较第一流表项和已成功下发到所述交换机上的第二流表项的优先级,获取优先级低于第一流表项的第二流表项;
 - 流表项通知单元,用于通知所述交换机删除所述第二流表项,并在交换机删除成功后,通知所述流表项下发单元重新向所述交换机下发所述第一流表项。
7. 如权利要求 6 所述的装置,其特征在于,所述表项优先级设置单元包括:

优先级列表构建子单元,用于预先构建业务优先级列表;

业务优先级查找子单元,用于根据所述第一流表项的请求报文中携带的业务标识信息,查找所述业务优先级列表;

业务优先级获取子单元,用于根据查找结果,获取所述请求报文对应的业务优先级参数;

表项优先级设置子单元,用于根据所述业务优先级参数,设置对应的所述第一表项优先级参数。

8. 如权利要求 7 所述的装置,其特征在于:

所述业务优先级查找子单元具体用于根据所述请求报文中携带的 IP 地址和 MAC 地址,查询预设的第一业务优先级列表,若所述第一业务优先级列表中无匹配表项,则根据所述请求报文携带的优先级字段信息,查询预设的第二业务优先级列表。

9. 如权利要求 6 所述的装置,其特征在于,所述流表项获取单元包括:

表项优先级比较子单元,用于将所述第一表项优先级参数与其保存的已成功下发到所述交换机上的流表项的表项优先级参数进行比较;

第二流表项选择子单元,用于根据比较结果,从表项优先级低于所述第一表项优先级的流表项中,选择优先级最低的流表项作为所述第二流表项。

10. 如权利要求 9 所述的装置,其特征在于:

所述第二流表项选择子单元,具体用于若存在多个所述优先级最低的流表项,则从所述多个优先级最低的流表项中选择报文匹配最少的流表项作为所述第二流表项。

流表下发方法及装置

技术领域

[0001] 本发明涉及网络通讯技术领域,尤其涉及一种 SDN(Software Defined Network,软件定义网络)网络中流表下发方法及装置。

背景技术

[0002] SDN 网络是一种新型网络创新架构,其核心思想是将网络设备的控制层面与转发层面分离,以实现网络流量的灵活控制,为核心网络及相关应用提供良好的平台。在 SDN 网络中,SDN 控制器搜集网络拓扑信息,计算转发路径,并将指导业务流量转发的流表下发到转发路径中的各个交换机上。交换机接收到业务报文后,按照 SDN 控制器下发的流表转发。

[0003] 由于交换机上可承载流表的硬件资源有限,而现有技术中对需要下发的流表采用尽量下发的方式,若交换机上有空闲的硬件资源,则下发成功;若没有空闲的硬件资源,则下发失败。但是,如果下发失败的流表项对应业务的业务优先级较高,则该业务由于无法通过流表进行转发,从而无法保证 SDN 网络中用户需求的高优先级业务的传输质量。

发明内容

[0004] 有鉴于此,本发明提供了一种流表下发方法,应用于软件定义网络 SDN 控制器,该方法包括:

[0005] 在接收到交换机上送的下发第一流表项的请求报文后,为所述第一流表项设置对应的第一表项优先级参数;

[0006] 向所述交换机下发所述第一流表项;

[0007] 在接收到所述交换机发送的所述第一流表项的下发流表溢出失败消息时,比较第一流表项和已成功下发到所述交换机上的第二流表项的优先级,获取优先级低于第一流表项的第二流表项;

[0008] 通知所述交换机删除所述第二流表项后,重新向所述交换机下发所述第一流表项。

[0009] 本发明还提供了一种流表下发装置,应用于软件定义网络 SDN 控制器,该装置包括:

[0010] 表项优先级设置单元,用于在接收到交换机上送的下发第一流表项的请求报文后,为所述第一流表项设置对应的第一表项优先级参数;

[0011] 流表项下发单元,用于向所述交换机下发所述第一流表项;

[0012] 流表项获取单元,用于在接收到所述交换机发送的所述第一流表项的下发流表溢出失败消息时,比较第一流表项和已成功下发到所述交换机上的第二流表项的优先级,获取优先级低于第一流表项的第二流表项;

[0013] 流表项通知单元,用于通知所述交换机删除所述第二流表项,并在交换机删除成功后,通知流表项下发单元重新向所述交换机下发所述第一流表项。

[0014] 在交换机硬件资源用尽的情况下,本发明通过设置业务优先级,删除低优先级业务,保证高优先级业务的正常转发,提高 SDN 网络的业务传输质量。

附图说明

- [0015] 图 1 是本发明一种实施方式中 SDN 网络示意图。
 [0016] 图 2 是本发明一种实施方式中流表下发方法的流程图。
 [0017] 图 3 是本发明一种实施方式中流表下发装置的基础硬件示意图。
 [0018] 图 4 是本发明一种实施方式中流表下发装置的结构示意图。
 [0019] 图 5 是本发明一种实施方式中表项优先级设置单元的结构示意图。
 [0020] 图 6 是本发明一种实施方式中第二流表项获取单元的结构示意图。

具体实施方式

[0021] 为使本发明的目的、技术方案及优点更加清楚明白,以下参照附图对本发明实施例作进一步地详细说明。

[0022] 图 1 所示为典型的 SDN 网络示意图。该 SDN 网络包括 SDN 控制器、交换机以及 PC 机,SDN 控制器与交换机之间通过 SDN 网络协议(例如:Openflow 协议)进行交互。SDN 控制器能够获取所述 SDN 网络的网络拓扑信息,当有访问请求时,根据搜集到的网络拓扑信息计算访问路径,并根据该路径生成对应的流表项下发到访问路径中的交换机上,交换机根据下发的流表项进行转发,实现对用户业务的处理。

[0023] 如图 1 所示,假设 PC1 的 IP 地址为 IP1,MAC 地址为 MAC1;PC2 的 IP 地址为 IP2,MAC 地址为 MAC2;PC3 的 IP 地址为 IP3,MAC 地址为 MAC3;PC4 的 IP 地址为 IP4,MAC 地址为 MAC4。PC1 与 PC2 在同一网段 subnet1,PC3 与 PC4 在同一网段 subnet2,子网掩码均为 Mask0。后续介绍中均以该假设为前提,不再重复说明。

[0024] 例如:当 PC1 访问 PC3 时,SDN 控制器根据网络拓扑信息计算的访问路径为 port3 → port1 → port2 → port5,根据该访问路径分别向访问路径中的 SW1 和 SW2 下发流表项。

[0025]

交换机	源 IP 地址	目的 IP	出端口
SW1	IP1	IP3	port1
SW2	IP1	IP3	port5

[0026] 表 1

[0027] 表 1 为 SDN 控制器根据计算出的访问路径生成的流表项,其中,第一条流表项下发给 SW1,第二条流表项下发给 SW2。交换机在接收到业务报文后,与其内部保存的 SDN 控制器下发的流表项进行匹配。例如,当交换机 SW1 接收到 PC1 访问 PC3 的业务报文时,根据该业务报文的源 IP 地址和目的 IP 地址进行流表项匹配,找到该业务报文的输出端口为 port1,则将该业务报文通过 SW1 的 port1 端口转发出去;同理,SW2 从 port5 端口将业务报文转发给 PC3,完成 PC1 到 PC3 的访问过程。

[0028] 然而,随着 SDN 网络的不断扩大,互访需求的逐步增多,下发到交换机上的流表项也会不断增加,但是交换机上的硬件资源有限,不能无限制接收下发的流表项。当交换机的硬件资源用尽时,即使有高优先级的业务需求,也无法立即向交换机下发,只能借由 SDN 控制器通过软件进行业务转发,或等待交换机中现有流表项老化,再进行新的流表项下发。前者会增加 SDN 控制器的负担,影响 SDN 控制器性能;后者无法保证高优先级业务的及时处理,降低了 SDN 网络的业务传输质量。

[0029] 针对上述问题,本发明实施例提出了一种流表下发方法,该方法应用于 SDN 控制器。SDN 控制器为每一个业务配置对应的业务优先级,当高优先级业务出现时,删除低优先级业务,保证高优先级业务的传输质量。

[0030] 请参见图 2,该方法包括:

[0031] 步骤 201,在接收到交换机上送的下发第一流表项的请求报文后,为所述第一流表项设置对应的第一表项优先级参数。

[0032] 交换机在接收到用户发送的业务报文时,首先查询本地已保存的 SDN 控制器下发的所有的流表项,若无与该业务报文匹配的流表项,则向 SDN 控制器发送下发第一流表项的请求报文,该第一流表项用于表征对应业务报文在当前交换机上的转发出口。SDN 控制器在接收到下发第一流表项的请求报文后,根据网络拓扑信息计算生成对应的第一流表项。

[0033] 在生成第一流表项后,SDN 控制器根据第一流表项请求报文中携带的业务标识信息,查找预设的业务优先级列表,其中,业务标识信息可以是请求报文中携带的源、目的 IP 地址和 / 或源、目的 MAC 地址。

[0034] 业务优先级列表可根据用户需求预先设置,其构建方式如下:

[0035] 1、根据实际的业务需要,定义业务优先级范围,例如:设定优先级可配置范围为 0 ~ 7,数值越大对应的优先级越高。

[0036] 2、SDN 控制器提供开放接口给用户,由用户自定义配置业务优先级策略,并将其定义为第一业务优先级列表。

[0037] 例如:在一个示例性的实施例中,用户可根据业务流的 IP/MAC 等定义的业务优先级。

[0038]

序号	业务匹配条件	业务优先级
1	Source IP = IP2&Source MAC = MAC2&Destination IP = IP4	7
2	(Source IP = IP1&Source MAC = MAC1) Destination IP = IP1	6
3	Source IP = subnet2&Mask0&Destination IP = any	5

[0039] 表 2

[0040] 表 2 为第一业务优先级列表示例。在该表中,序号为 1 业务优先级表项表示将 PC2(源 IP 地址为 IP2,且源 MAC 地址为 MAC2)访问 PC4(目的 IP 地址为 IP4)的业务优先级设置为 7;序号为 2 的业务优先级表项表示将 PC1(源 IP 地址为 IP1,且源 MAC 地址为 MAC1,

或者目的 IP 地址为 IP1) 所有的业务优先级均设置为 6;序号为 3 的业务优先级表项表示将网段 subnet2 内所有主机 (PC3 和 PC4) 对外访问的业务优先级设置为 5。因此,当发生业务访问时,可通过业务报文中携带的 IP/MAC 地址信息,查询第一业务优先级列表,获取对应业务优先级。

[0041] 3、对于其他未在第一业务优先级列表中定义的业务,根据交换机上送的 packet in 用户报文中携带的 DSCP (Differentiated Services Code Point, 差分服务代码点) 值或者 802.1p 字段中的值来映射对应的业务优先级,在本申请中将其定义为第二业务优先级列表。

[0042]

DSCP	802.1p	业务优先级
0 ~ 7	0	0
8 ~ 15	1	1
16 ~ 23	2	2
24 ~ 31	3	3
32 ~ 39	4	4
40 ~ 47	5	5
48 ~ 55	6	6
56 ~ 63	7	7

[0043] 表 3

[0044] 表 3 为本申请一个例子中第二优先级列表的示例。其中, DSCP 为 IP 报文的优先级字段, 802.1p 表示以太网报文的优先级字段。用户可根据实际需要自定义优先级字段取值范围与业务优先级的对应关系。例如,若接收到的业务报文为 IP 报文,该 IP 报文的 DSCP 字段信息为 50,则通过查表(表 3)可知,该业务报文对应的业务优先级为 6。

[0045] SDN 控制器在接收到交换机上送的请求报文后,首先查询预设的第一业务优先级表,若存在匹配表项,则将该匹配表项对应的业务优先级作为该请求报文对应的业务优先级;若不存在匹配表项,则根据所述请求报文的优先级字段信息(DSCP 或 802.1p),查询预设的第二业务优先级表,将所述优先级字段信息对应的业务优先级作为该请求报文对应的业务优先级。将通过上述方式获得的业务优先级作为第一表项优先级参数。

[0046] 步骤 202,向所述交换机下发第一流表项。

[0047] SDN 控制器在生成对应业务的第一流表项后,将该第一流表项下发给交换机,指导交换机按照该第一流表项进行业务报文转发。

[0048] 步骤 203,在接收到所述交换机发送的所述第一流表项的下发流表溢出失败消息时,比较第一流表项和已成功下发到所述交换机上的第二流表项的优先级,获取优先级低

于第一流表项的第二流表项。

[0049] 交换机在接收到 SDN 控制器下发的第一流表项后,查询本地是否存在空闲的硬件资源,若存在,则保存该第一流表项,按照该第一流表项进行转发;若不存在,则向 SDN 控制器发送下发溢出失败消息。SDN 控制器接收到该下发失败消息后,从该消息中获知是由于流表溢出(交换机中保存的流表数量已达到上限)而导致的下发流表失败,则通过将下发失败的第一流表项与本地保存的已成功下发到同一交换机上的流表项的表项优先级进行比较,从已成功下发到同一交换机上的流表项中,找到表项优先级低于第一表项优先级的流表项,若存在多个表项优先级低于第一表项优先级的流表项,则找出表项优先级最低的流表项,若存在多个表项优先级最低的流表项,则 SDN 控制器从交换机获取统计的报文匹配次数,选择业务报文匹配最少的流表项作为第二流表项,后续通知交换机删除该流表项。通过删除利用率最低的流表项,达到最优的优先级处理效果。

[0050] 步骤 204,通知所述交换机删除所述第二流表项,重新向所述交换机下发所述第一流表项。

[0051] SDN 控制器在确认了需要删除的第二流表项后,向交换机下发删除命令,并重新下发第一流表项。交换机根据 SDN 控制器下发的删除命令,删除对应的流表项,释放硬件资源,当再次接收到第一流表项时,即可将该第一流表项保存在本地,并按照该第一流表项进行对应业务报文的转发。

[0052] 现仍以图 1 所示 SDN 网络为例,介绍本方案中流表下发的处理过程。

[0053] 假设,PC2 请求访问 PC4,SW1 查询本地无对应的流表项,则向 SDN 控制器请求下发对应的流表项。SDN 控制器根据网络拓扑信息计算出应该下发给 SW1 的流表项,并查询第一业务优先级列表,如表 4 所示。

[0054]

序号	业务匹配条件	业务优先级
1	Source IP = IP1&Source MAC = MAC1&Destination IP = IP3	1
2	Source IP = IP1&Source MAC = MAC1&Destination IP = IP4	1
3	Source IP = IP2&Source MAC = MAC2&Destination IP = IP3	5
4	Source IP = IP2&Source MAC = MAC2&Destination IP = IP4	7

[0055] 表 4

[0056] 根据源 IP 为 IP2,源 MAC 为 MAC2,以及目的 IP 为 IP4,查询表 4,可知该业务与第 4 条业务优先级表项匹配,则对应的业务优先级为 7,将该业务优先级作为流表项的表项优先级。SDN 控制器针对 PC2 访问 PC4 的业务下发给 SW1 的流表项以及为该流表项配置的表项优先级如表 5 所示。

[0057]

交换机	源 IP 地址	目的 IP 地址	出端口	表项优先级
SW1	IP2	IP4	port1	7

[0058] 表 5

[0059] SW1 接收到该流表项后,发现本地无硬件资源可用,则向 SDN 控制器发送流表项下发失败的消息。SDN 控制器接收到该下发失败消息,并从该消息中确认是由于 SW1 硬件资源用尽导致的流表项下发失败,则将本次下发失败的流表项的表项优先级与本地保存的已成功下发到 SW1 的流表项的表项优先级进行比较。假设已成功下发到 SW1 的流表项以及对应表项优先级如表 6 所示。

[0060]

交换机	源 IP 地址	目的 IP	出端口	表项优先级
SW1	IP1	IP3	port1	1
SW1	IP1	IP4	port1	1
SW1	IP2	IP3	port1	5

[0061] 表 6

[0062] 查询表 6,找到表项优先级低于 7 的流表项。从表 6 中可以看到,所有流表项对应的表项优先级均低于 7,则继续查找表项优先级最低的流表项,第一条流表项和第二条流表项的表项优先级均为 1,优先级最低。在同时存在多个最低优先级的流表项时,通过确认哪一个流表项的报文匹配次数最少,即使用频率最低,则删除该匹配次数最少的流表项。假设第一条流表项的匹配次数最少,则通知 SW1 删除该流表项,重新下发上次下发失败的流表项给 SW1,同理,SW2 上的流表项下发方法相同,从而保证 PC2 访问 PC4 的高优先级业务的服务质量。

[0063] 与上述方法实施例对应的是,本发明实施例还提供一种流表下发装置。以下以软件实现为例,但是本发明并不排除诸如硬件或者逻辑器件等其他实现方式。请参考图 3,该装置运行的硬件环境包括 CPU、内存、非易失性存储器以及其他硬件。该装置作为一个逻辑层面的虚拟装置,其通过 CPU 来运行。请参考图 4,该流表下发装置 40 包括:表项优先级设置单元 401、流表项下发单元 402、流表项获取单元 403 以及流表项通知单元 404,其中:

[0064] 表项优先级设置单元 401,用于在接收到交换机上送的下发第一流表项的请求报文后,为所述第一流表项设置对应的第一表项优先级参数;

[0065] 流表项下发单元 402,用于向所述交换机下发所述第一流表项;

[0066] 流表项获取单元 403,用于在接收到所述交换机发送的所述第一流表项的下发流表溢出失败消息时,比较第一流表项和已成功下发到交换机上第二流表项的优先级,获取优先级低于第一表项优先级的第二流表项;

[0067] 流表项通知单元 404,用于通知所述交换机删除所述第二流表项,并在交换机删除成功后,通知所述流表项下发单元 402 重新向所述交换机下发所述第一流表项。

[0068] 请参考图 5,表项优先级设置单元 401 可以具体包括:优先级列表构建子单元 4011、业务优先级查找子单元 4012、业务优先级获取子单元 4013 以及表项优先级设置子单元 4014,其中:

[0069] 优先级列表构建子单元 4011,用于预先构建业务优先级列表;

[0070] 业务优先级查找子单元 4012,用于根据所述第一流表项的请求报文中携带的业务

标识信息,查找所述业务优先级列表;

[0071] 业务优先级获取子单元 4013,用于根据查找结果,获取所述请求报文对应的业务优先级参数;

[0072] 表项优先级设置子单元 4014,用于根据所述业务优先级参数,设置对应的所述第一表项优先级参数。

[0073] 进一步地,业务优先级查找子单元 4012 具体用于根据所述请求报文中携带的 IP 地址和 / 或 MAC 地址,查询预设的第一业务优先级列表,若所述第一业务优先级列表中无匹配表项,则根据所述请求报文携带的优先级字段信息,查询预设的第二业务优先级列表。

[0074] 请参考图 6,第二流表项获取单元 403 可以具体包括:表项优先级比较子单元 4031 和第二流表项选择子单元 4032,其中:

[0075] 表项优先级比较子单元 4031,用于将所述第一表项优先级参数与其保存的已成功下发到所述交换机上的流表项的表项优先级参数进行比较;

[0076] 第二流表项选择子单元 4032,用于根据比较结果,从表项优先级低于所述第一表项优先级的流表项中,选择优先级最低的流表项作为所述第二流表项。

[0077] 进一步地,第二流表项选择子单元 4032,可以具体用于若存在多个所述优先级最低的流表项,则从所述多个优先级最低的流表项中选择报文匹配最少的流表项作为所述第二流表项。

[0078] 本发明实施例提供的流表下发装置可以保证高优先级业务的正常转发,其具体实现过程可参见上述方法实施例的说明,在此不再赘述。

[0079] 从以上方法和装置的实施方式中可以看出,在交换机硬件资源用尽的情况下,通过设置业务优先级,删除低优先级业务,保证高优先级业务的正常转发,提高 SDN 网络的业务传输质量。

[0080] 以上所述仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明保护的范围之内。

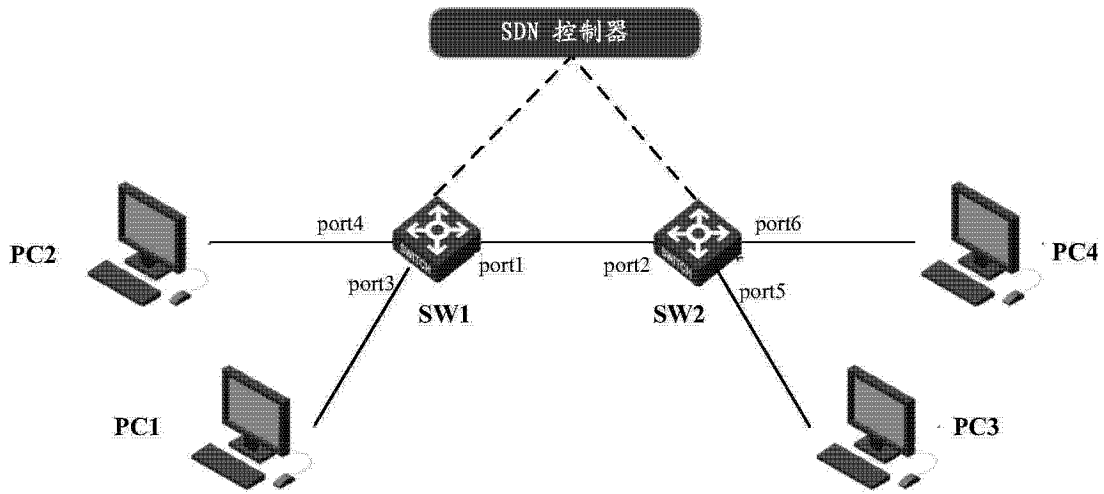


图 1

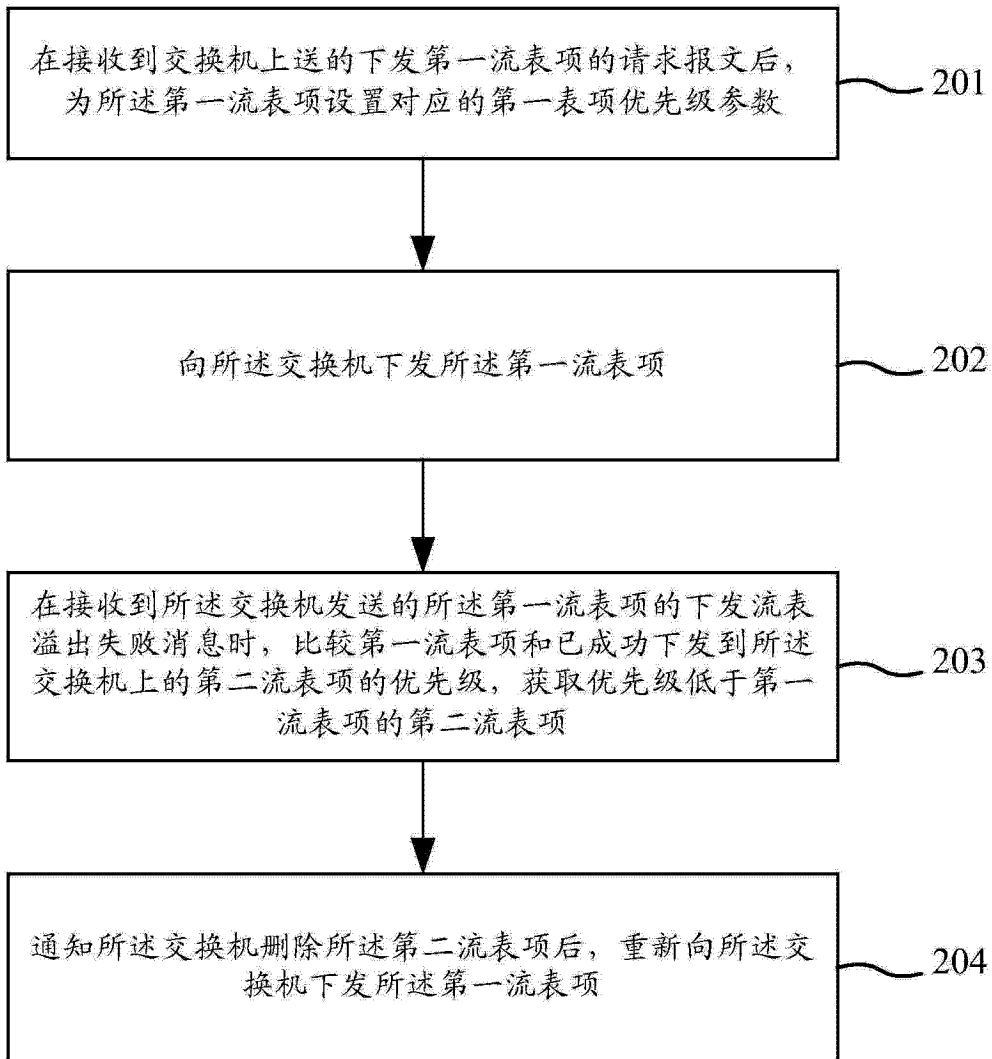


图 2

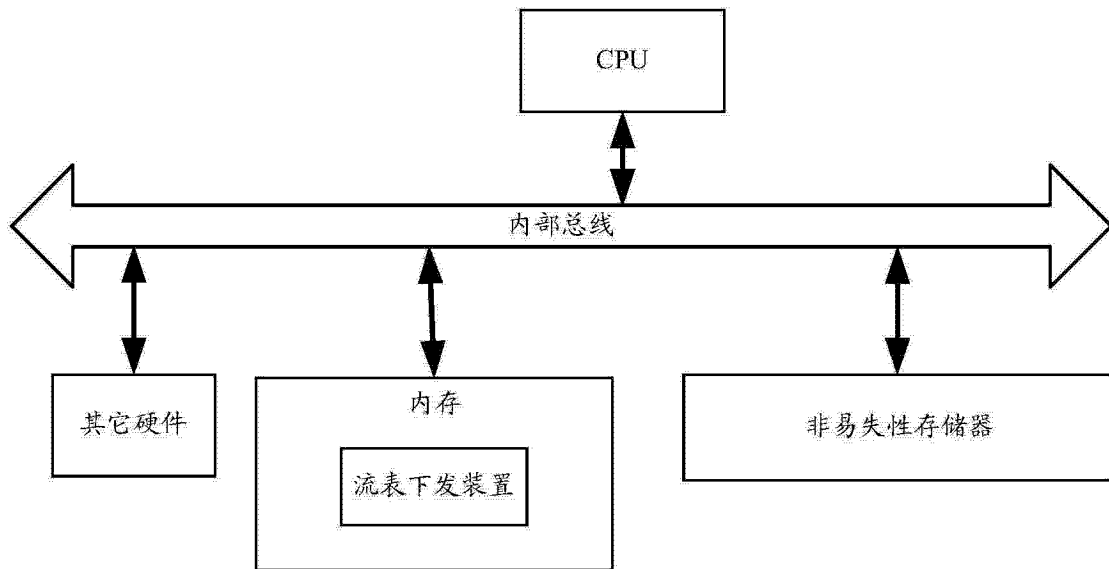


图 3

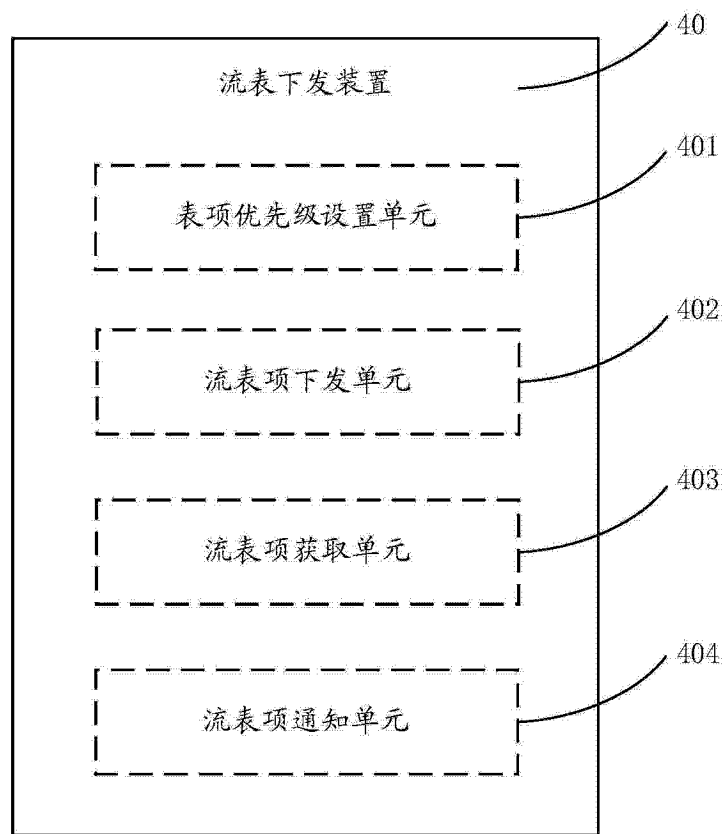


图 4

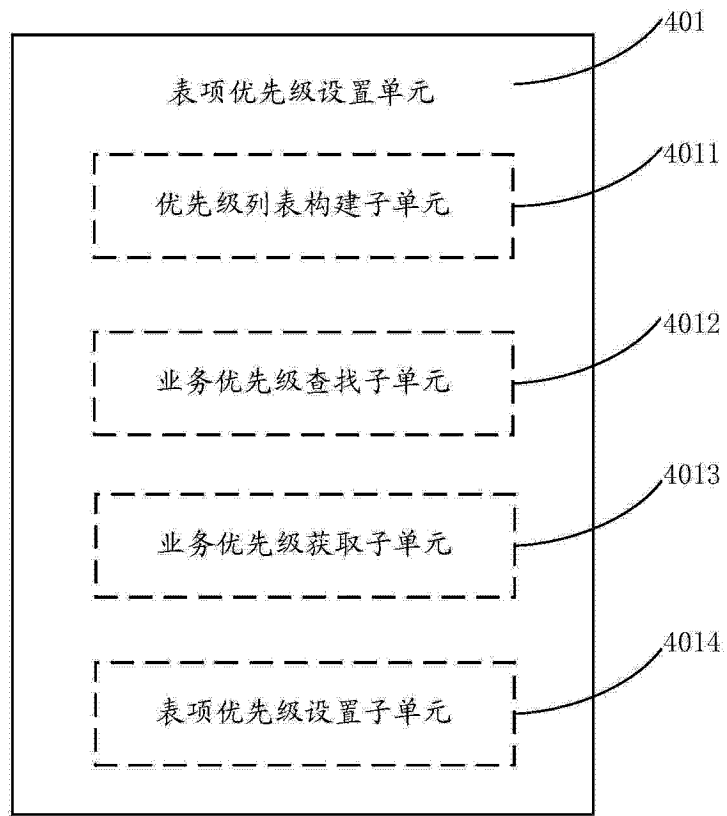


图 5

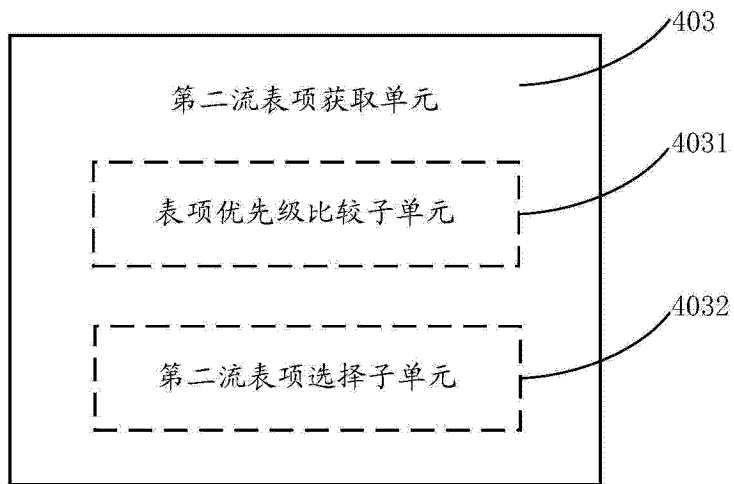


图 6