

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H04L 12/56

H04L 12/24 H04Q 3/545



[12] 发明专利申请公开说明书

[21] 申请号 03100352.4

[43] 公开日 2004年8月4日

[11] 公开号 CN 1518296A

[22] 申请日 2003.1.13 [21] 申请号 03100352.4

[71] 申请人 华为技术有限公司

地址 517057 广东省深圳市科技园科发路华为用户服务中心大厦知识产权部

[72] 发明人 安 雁 薛国锋

[74] 专利代理机构 北京集佳知识产权代理有限公司

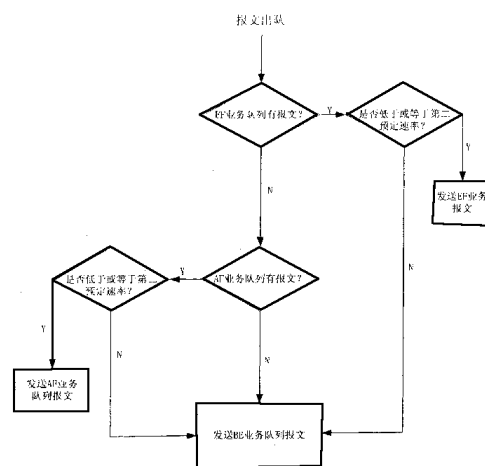
代理人 王学强

权利要求书 3 页 说明书 11 页 附图 3 页

[54] 发明名称 一种支持多业务的综合队列调度的实现方法

[57] 摘要

本发明涉及一种支持多业务的综合队列调度方法，包括步骤：当第一预定义优先报文入队时对其不同类的报文进入第一优先发送组相应队列进行第一预定速率限制；如进队速率低于或等于第一预定速率则进入所述类预留的第一优先发送组相应队列，否则丢弃；当第一预定义优先报文或第二预定义优先报文出队时如检测到默认队列有报文则对所述第一或第二预定义报文出队发送进行第二预定速率限制；如第一或第二预定义优先报文出队发送速率低于或等于所述的第二预定速率则可出队发送；否则发送第三预定义报文。本发明通过绝对优先发送和令牌桶速率限制的方法，保证优先业务报文发送质量的同时，解决了某些业务报文由于没有可用带宽而永远无法得到调度的问题。



1、一种支持多业务的综合队列调度实现方法，所述方法包括业务报文进队列步骤和出队列步骤，其特征在于，所述业务报文进队列步骤包括：

11)当第一预定义优先报文入队时对其不同类的报文进入第一优先发送组相应队列进行第一预定速率限制；

12)第二预定义优先报文进入第二优先发送组队列；

13)第三预定义报文进入默认队列；

其中如所述的第一预定义优先报文进队速率低于或等于该第一预定速率则该第一预定义优先报文进入该报文相应类所预留的第一优先发送组相应队列，否则丢弃；

业务报文出队列步骤包括：

21)首先发送第一优先发送组队列中的第一预定义优先报文；

22)待第一优先发送组队列空后再发送第二优先发送组队列中第二预定义优先报文；

23)最后发送默认队列中的第三预定义报文；

24)实时检测默认队列是否有所述第三预定义报文。

2、根据权利要求1所述的一种支持多业务的综合队列调度实现方法，其特征在于还包括：若所述步骤24)检测到该默认队列有该第三预定义报文，则当所述第一预定义优先报文或第二预定义优先报文明出队发送时对所述第一或第二预定义优先报文明出队发送的速率进行第二预定速率限制，如所述第一或第二预定义优先报文明出队速率低于或等于所述的第二预定速率时则可以出队发送；

如所述第一或第二预定义优先报文明出队速率大于该第二预定速率

则发送默认队列中的第三预定义报文；

若步骤 24) 检测到该默认队列无第三预定义报文则不对第一或第二优先发送报文出队进行第二预定速率限制, 重复执行步骤 21)~22)。

3、根据权利要求 1 所述的一种支持多业务的综合队列调度实现方法, 其特征在于所述的第一预定义优先报文为 EF (Expediated Forwarding) 业务报文。

4、根据权利要求 1 所述的一种支持多业务的综合队列调度实现方法, 其特征在于所述第二预定义优先报文为 AF (Assured Forwarding) 业务报文。

5、根据权利要求 1 所述的一种支持多业务的综合队列调度实现方法, 其特征在于所述第三预定义报文为 BE (Best-Effort Forwarding) 业务报文。

6、根据权利要求 3、4 或 5 所述的一种支持多业务的综合队列调度实现方法, 其特征在于所述的 EF 业务报文进入第一优先发送组队列首先发送, AF 业务报文进入第二优先发送组队列待第一优先发送组队列空再发送, BE 业务报文进入默认队列最后发送。

7、根据权利要求 1 所述的一种支持多业务的综合队列调度实现方法, 其特征在于所述的第一预定速率限制和第二预定速率限制均采用令牌桶算法进行。

8、根据权利要求 7 所述的一种支持多业务的综合队列调度实现方法, 其特征在于所述的第一预定速率取第一令牌桶平均速率, 该第一令牌桶平均速率采用用户为所述类配置的带宽, 突发量取一个最大报文的长度; 所述第二预定速率取第二令牌桶平均速率, 该第二令牌桶平均速率采用在接口上应用的各个业务的实际带宽和, 最大不超过该

接口的最大预留带宽，突发量取一个报文的最大长度。

9、根据权利要求6所述的一种支持多业务的综合队列调度实现方法，其特征在于对AF业务报文进行队列操作时包括确定结束势能的步骤，以便当EF业务队列空时选择AF业务队列中最小结束势能的报文出队发送，包括下面的步骤：

确定当前系统势能；

取该报文入队前的系统势能和队列势能的较大值作为所述报文的起始势能；

根据报文长度与所配置带宽的比值再与起始势能的和作为该报文的结束势能。

10、根据权利要求6所述的一种支持多业务的综合队列调度实现方法，其特征在于对BE业务报文进队列时根据报文的源IP地址、目的IP地址、源端口、目的端口、TOS（服务类型）、协议字特征区分一条流，散列到BE的各个队列，出队列时采用配额轮询算法调度发送，其中所述的配额是最大预留带宽在分配给所述的EF业务报文和AF业务报文所需带宽后剩余的可分配带宽。

一种支持多业务的综合队列调度的实现方法

技术领域

本发明涉及数据通信领域，尤其涉及一种支持多业务的综合队列调度方法。

背景技术

在因特网分组交换的复杂环境下，网络拥塞极为常见。拥塞使流量不能及时获得资源，是造成服务性能下降的源头，拥塞有可能会引发以下的负面影响：拥塞增加了报文传输的延迟和延迟抖动，过高的延迟会引起报文重传；拥塞使网络的有效吞吐率降低，造成网络资源的损害；拥塞加剧会耗费大量的网络资源（特别是存贮资源），不合理的资源分配甚至可能导致系统陷入资源死锁而崩溃。然而在分组交换以及多用户业务并存的复杂环境下，拥塞又是常见的，网络发生拥塞时必须对其进行管理和控制，常见方法是使用队列技术，亦可称为排队策略。拥塞管理就是将所有要从一个接口发出的报文进行分类送入不同的队列，然后按照不同的策略进行处理；这些策略用来处理对带宽的要求超出了网络能提供的总带宽的情况。拥塞管理在拥塞发生之前并不动作，拥塞时它控制进入网络的报文流量，以让某种网络流量有相对其他流量较高的优先权。下面是几种常见的排队策略：

FIFO——先进先出队列。不对报文进行分类，当报文进入接口的速度大于接口能发送的速度时，FIFO按报文到达接口的先后顺序让报文进入队列，同时，在队列的出口报文按入队的顺序出队，先进的报

文将先出队，后进的报文将后出队。

PQ——优先队列。对报文进行分类，按报文的类别将报文送入相应的队列。分类时属于较高优先级队列的报文将会得到优先发送，而较低优先级的报文将会在发生拥塞时被较高优先级的报文抢先。

CQ——定制队列。对报文进行分类，按报文的类别将报文进入相应的队列。在报文出队的时候，按定义的带宽比例分别从队列中取一定量的报文在接口上发送出去。

WFQ——加权公平队列。对报文按流进行分类，每个流被分配到一个队列。出队时，按流的优先级分配每个流应占有出口的带宽。优先级的数值越小，所得带宽越少。反之，所得带宽越多。

综上，FIFO 队列技术对各种业务一视同仁，网络拥塞时一同丢弃；PQ 队列技术虽然从某种意义上保证了某些业务的优先，但是一方面由于只有 4 个队列，承载的业务太少，另一方面如果优先的业务持续发送，可能导致其他业务永远无法得到调度；CQ 队列技术有 16 个队列，循环调度各个队列，这样又无法保证语音等优先的业务；WFQ 相比前几种队列技术，有所改进，队列数目相对增加，优先级越高，权重越大，队列根据权重进行加权公平调度，但是和 CQ 队列技术一样，无法真正保证多种业务的要求。因此，现有的各种队列技术很难满足日益广泛的多种业务并存、区分服务的要求。

发明内容

本发明的目的是提供一种实现支持多业务的综合队列调度的方法，更确切的本发明根据网络上的各种业务的特点，提出了一种多业务并存、区分服务的综合队列调度的方法。

根据本发明的方法，一种支持多业务的综合队列调度实现方法，

所述方法包括业务报文进队列步骤和出队列步骤，该业务报文进队列步骤包括：

11)当第一预定义优先报文入队时对其不同类的报文进入第一优先发送组相应队列进行第一预定速率限制；

12)第二预定义优先报文进入第二优先发送组队列；

13)第三预定义报文进入默认队列；

其中如所述的第一预定义优先报文进队速率低于或等于该第一预定速率则该第一预定义优先报文进入该报文相应类所预留的第一优先发送组相应队列，否则丢弃；

业务报文出队列步骤包括：

21)首先发送第一优先发送组队列中的第一预定义优先报文；

22)待第一优先发送组队列空后再发送第二优先发送组队列中第二预定义优先报文；

23)最后发送默认队列中的第三预定义报文；

24)实时检测默认队列是否有所述第三预定义报文；

在上述步骤中，若步骤 24)检测到该默认队列有该第三预定义报文则当所述第一预定义优先报文或第二预定义优先报文出队发送时对所述第一或第二预定义优先报文出队发送的速率进行第二预定速率限制，如所述第一或第二预定义优先报文出队发送速率低于或等于所述的第二预定速率则可以出队发送；

如所述第一或第二预定义优先报文出队速率大于该第二预定速率则发送默认队列中的第三预定义报文；

若步骤 24)检测到该默认队列无第三预定义报文则不对第一或第二优先发送报文出队进行第二预定速率限制，重复执行步骤 21)~22)。

具体的所述的第一预定义优先报文为 EF (Expediated Forwarding) 业务报文。

具体的所述第二预定义优先报文为 AF (Assured Forwarding) 业务报文。

具体的所述第三预定义报文为 BE (Best-Effort Forwarding) 业务报文。

进一步所述的 EF 业务报文进入第一优先发送组队列首先发送，AF 业务报文进入第二优先发送组队列待第一优先发送组队列空再发送，BE 业务报文进入默认队列最后发送。

其中所述的第一预定速率限制和第二预定速率限制均采用令牌桶速率算法进行速率限制，所述的第一预定速率取第一令牌桶平均速率，该第一令牌桶平均速率采用用户为所述类配置的带宽，突发量取一个最大报文的长度，所述第二预定速率取第二令牌桶平均速率，该第二令牌桶平均速率采用在接口上应用的各个业务的实际带宽和，最大不超过该接口最大预留带宽，突发量取一个最大报文的长度。

另外对 AF 业务报文进行队列操作时包括确定结束势能的步骤，以每当 EF 业务队列空时选择 AF 业务队列中最小结束势能的报文出队发送，包括下面的步骤：

确定当前系统势能；

取该报文入队前的系统势能和队列势能的较大值作为所述报文的起始势能；

根据报文长度与所配置带宽的比值再与起始势能的和作为该报文的结束势能。

具体的对 BE 业务报文进队列时根据报文的源 IP 地址、目的 IP

地址、源端口、目的端口、TOS（服务类型）、协议字等特征区分一条流，散列到BE的各个队列，出队列时采用配额轮询算法调度发送，其中所述的配额是最大预留带宽在分配给所述的EF业务报文和AF业务报文所需带宽后剩余的可分配带宽。

本发明通过绝对优先发送和令牌桶速率限制的方式，既保证了高优先级报文的发送质量，同时又通过速率限制的方式避免了那些低优先级队列的报文由于没有可用的带宽而永远无法得到发送的问题，达到区分服务，同时保证多业务并存的问题，可以使报文的带宽、延迟、延迟抖动得到保证，降低网络维护费用。

附图说明

图1 描述了本发明多业务调度顺序的流程图；

图2 描述了本发明AF业务公平算法的示意图；

图3 描述了本发明多类业务绝对优先调度和令牌桶速率限制相结合的流程图；

具体实施方式

为了让本领域的技术人员更好的理解本发明的方法，下面结合附图进一步描述本发明的具体实施方式。

在本发明中我们预定义了三种不同服务类型的报文，分别是第一预定义优先报文和第二预定义优先报文以及第三预定义报文，分别代表不同业务特点的业务报文，其中，所述的第一预定义优先报文是需要最先发送的报文，所述的第二预定义报文是比第一预定义优先报文发送级别低，次发送的报文，所述的第三预定义报文则是最后发送的报文，对应的，所述的不同预定义报文分类进入不同的优先发送组队列发送，具体的第一预定义优先报文进入第一优先发送组队列最先发

送，第二预定义优先报文进入第二优先发送组队列待第一优先发送组队列空次发送，第三预定义报文则进入默认队列最后发送。我们知道目前网络上的业务大体可以划分为：EF 业务、AF 业务和 BE 业务。在本发明的具体实施方式中，我们所指的第一预定义优先报文为 EF 业务报文，第二预定优先报文为 AF 业务报文，第三预定义报文为 BE 业务报文，下面我们就以这几种业务为例具体描述本发明的具体实施方式，根据不同业务的业务特点，本发明中采用不同的队列组承载不同的业务的方法，同时在出队列发送时按照绝对优先发送的原则，保证优先级高的报文优先发送，另外，在本发明中侧重一种队列调度的方法，对报文分类进队可采用 CBWFQ(基于类的加权公平队列)等技术进行，在此不再细述，仅就本发明提出的队列调度方法进行详述。

下面请参看图 1 所示，根据各业务报文的特点，在本实施方式中业务报文的发送顺序是 EF 业务报文最先发送，AF 业务报文次发送，BE 业务报文最后发送，即只有在 EF 业务队列空后，才能发送 AF 业务队列的报文，AF 业务队列空后才发送 BE 业务报文，更进一步阐述本发明的方法，我们先来看 EF 业务队列的业务特点，EF 业务队列针对实时应用程序，主要包括二层协议报文、语音业务报文等。EF 业务报文要求具有严格的延时、抖动保障机制和时钟同步机制。而导致分组延迟和分组抖动的主要原因是排队延迟，排队延迟是因为队列非常长引起的，通常在网络拥塞点会出现这种情况。当分组最大到达速率小于离开速率，则基本上可以消除排队延迟。因此，该业务队列的关键问题是尽快将报文传送出去，最大限度的降低延迟和抖动。

基于以上分析，由于二层协议报文数量有限，而且担负着维护链路状态的重任，因此本发明中采用二层协议报文入队时不做任何限制，

出队时优先发送的原则；对于语音报文，其对时延和抖动也非常敏感，可采用入队时令牌桶速率限制，出队时绝对优先发送的原则。

下面再看AF业务特点，AF业务针对在拥塞发生时要求保证带宽的业务。因此，关键问题是确保带宽。

由于队列是按包进行发送的，因此无法理想的模拟GPS（广义处理器共享）系统。GPS的原理是将每个流都放在各自的逻辑队列中，然后一次为每个非空队列传输无穷小的数据量，每一轮都只传输无穷小的数据量，因此在任何有限的时间间隔内，所有的非空队列都会被访问，因此任何时候都是公平的。

在本实施方式中对AF业务队列我们寻找与GPS接近并且可以实现的模型。具体的我们引进势能的概念来跟踪系统的状态，每个连接（队列或者类）关联一个队列势能来记录接口拥塞情况下该队列得到的服务时间与尚需的服务时间之和；每个接口关联一个系统势能来记录该接口自拥塞以来服务的时间；每个包关联一个起始势能和一个结束势能，分别用来记录包到来的时间和包预计发出的时间。这些势能都是一个拥塞期间关于时间的广义递增函数。一个包入队时，通过时间戳，包长以及当前的系统势能就能唯一的确定一个值，这个值唯一的决定当前队列中报文的发送顺序，队列调度到AF业务队列时选取一个最小结束势能的报文进行发送。

请参看图2本发明AF业务公平算法的示意图所示，每个队列有自己的队列势能（ P_i ），每个报文有自己的开始势能（SP）和结束势能（TS），每个接口有自己的系统势能（ P_s ），同时我们定义 b_i 为队列 i 需要保证的带宽， b 为该接口的可用带宽， v 为非空的队列个数。 P_s 和 P_i 的初始值都为0，在本发明中势能的单位为纳秒。

当一个接口发生拥塞时，所有的报文在发送的时候都需要经过队列调度，简单的说，需要进行入队列操作和出队列操作。以第 i 个队列到来的第 k 个报文为例，此时该报文的SP为该报文入队前的 P_s 和 P_i 的较大值，计算该报文的TS等于该报文长度与所配置带宽的比值再与SP的和，该报文入队，相应的 P_i 也增加为新的TS值，这样就保证了接口队列中的每个报文都有自己的TS值，而该TS值是报文出队时衡量的唯一标准。当物理层来一个中断准备发送一个报文的时候，此时由于接口拥塞，从队列中选择一个报文发出，如果调度到AF队列，AF队列一般由若干队列组成，AF从各个队列中挑选一个最小TS值的报文发出，请再参看图2所示，报文中标注的数字表示该报文的调度发送次序，其中数字1表示该报文的TS值最小，最先发送，2表示该报文的TS值次之，第二发送，当然，由于AF队列中每次进一个报文或调度发送一个报文，AF队列中的所有报文的TS值都要重新计算，所以发送次序也是符合公平的，另外，由于报文以系统的可用带宽速率发出，因此，需要更新系统势能，这样才能保证TS值小于 P_s 的报文都被发出，最大限度的模拟GPS系统的工作。

AF业务队列出队列的优先级仅次于EF业务队列的报文，直到EF队列没有报文才能调度到AF业务队列。

由于AF业务各个队列之间的带宽是可以自适应的，某一时刻AF队列中某个队列没有报文时，AF队列中的其他队列可以按比例占用该带宽。

本发明中AF在出队时考虑了报文入队的时间，报文的大小，最大限度的保证了报文出队的公平性，降低了时延。

最后再看BE业务，BE业务属于尽力传送的数据业务，如上网业务、

FTP业务。一般，BE业务队列针对不需要保证延时和带宽的业务，因此，该类业务的关键问题是保证能够发得出去，同时保证BE业务队列间的相对公平性。

本发明中BE业务队列由多个队列组成，在进队列前根据报文的源IP地址、目的IP地址、源端口、目的端口、TOS（服务类型）、协议字等特征区分一条流，散列到BE的各个队列。出队列时通过轮询算法尽量保证公平性。

由于BE业务属于尽力传送的业务，BE业务队列出队时采用配额轮询算法，当调度到BE业务队列时，循环从各个队列发送配额，配额不足者等待下一次调度，所述的配额是最大预留带宽在分配给所述的EF业务报文和AF业务报文所需带宽后剩余的可分配带宽，具体的如最大的接口带宽为总带宽的75%，则当分配给EF业务报文和AF业务的占60%，则有15%的配额可分配给BE业务报文。

在本实施方式中，BE业务队列与EF业务队列和AF业务队列相比是优先级最低的队列，最后才被调度。

上述在队列调度过程中，EF业务队列优先、AF业务队列次之，BE业务队列最后。这样会严格保证高优先级队列的带宽和延迟，同时也可能造成某些情况下BE业务队列永远无法得到调度。然而从各种业务的特征考虑，AF业务对延时并不敏感。本发明采用了绝对优先调度和令牌桶速率限制相结合的方式解决该问题。EF业务在入队时以及EF业务队列和AF业务队列出队时都要经过令牌桶速率限制。

本实施方式中，当EF业务进队进入第一优先发送组队列时，直接对其不同类的报文进入相应类所预留的第一优先发送组相应队列进行第一预定速率限制，其中所述的第一预定速率取第一令牌桶平均速率，

该第一令牌桶平均速率可采用用户为该类配置的带宽，突发量取一个最大报文的长度。如果低于或等于该第一令牌桶速率，则进入相应类所预留的第一优先发送组相应队列，否则丢弃。

再请参看图3本发明多类业务绝对优先调度和令牌桶速率限制相结合的流程图，报文出队发送时具体步骤如下：

步骤1，判断EF业务队列即所述的第一优先发送组队列是否有报文，如有所述的EF业务报文，则对该EF业务报文出队发送进行第二预定速率限制，其中所述的第二预定速率取另一个令牌桶速率，为了区别这里记为第二令牌桶速率，在本实施方式中，第二令牌桶平均速率采用在接口上应用的各个业务的实际带宽和，最大不超过接口的最大预留带宽（由于BE业务队列，二层协议队列报文、物理层开销、ATM分固定信元所占的开销需要占用一部分带宽，接口实际带宽和这部分带宽的差值为最大预留带宽），突发量仍然为一个最大报文的长度。这样如所述的EF业务出队发送速率低于或等于所述的速率限制则出队发送，否则执行步骤3；

步骤2，判断AF业务队列是否有报文，即所述的第二优先发送组队列是否有所述的AF业务报文，如有则对其报文出队发送进行第二预定速率限制，具体的所述的第二预定速率仍为第二令牌桶的平均速率即采用在接口上应用的各个业务的实际带宽和，最大不超过接口最大预留带宽，当所述的AF业务报文出队发送速率低于或等于所述的第二预定速率则继续发送，否则执行步骤3；

步骤3，调度默认队列中的BE业务报文出队发送；

步骤4，实时检测默认队列是否有BE业务报文，如果有，则按步骤1或步骤2对第一预定义优先报文或第二预定义优先报文出队发送进行

第二预定速率限制，否则并不进行第二预定速率限制，按照绝对优先发送的原则发送所述的EF、AF业务报文。

不难理解，经过本发明的绝对优先发送和令牌桶速率限制，既保证了EF、AF等优先级高报文的发送质量又防止BE业务由于没有可用的带宽而永远无法得到调度的问题。

以上所述，仅为本发明的一可行实施例而已，非因此即局限本发明的权利范围，凡运用本发明说明书及附图内容所为的等效变化，均理同包含于本发明的权利要求范围内。

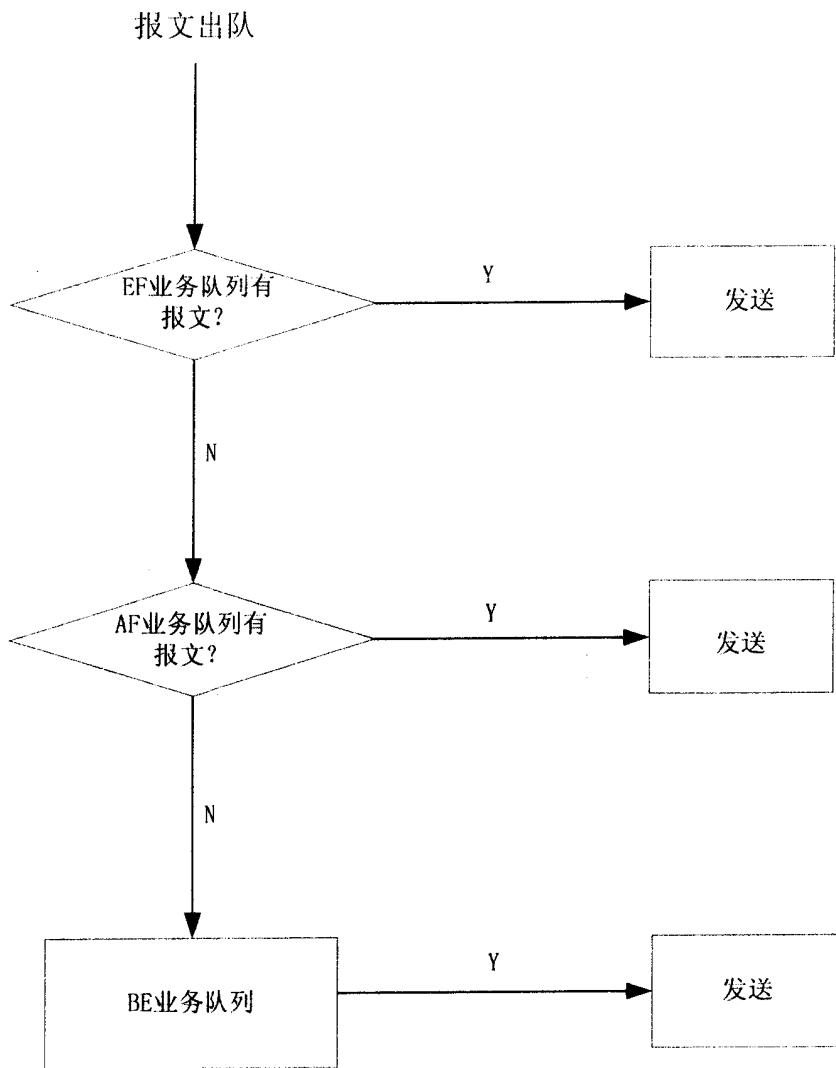


图1

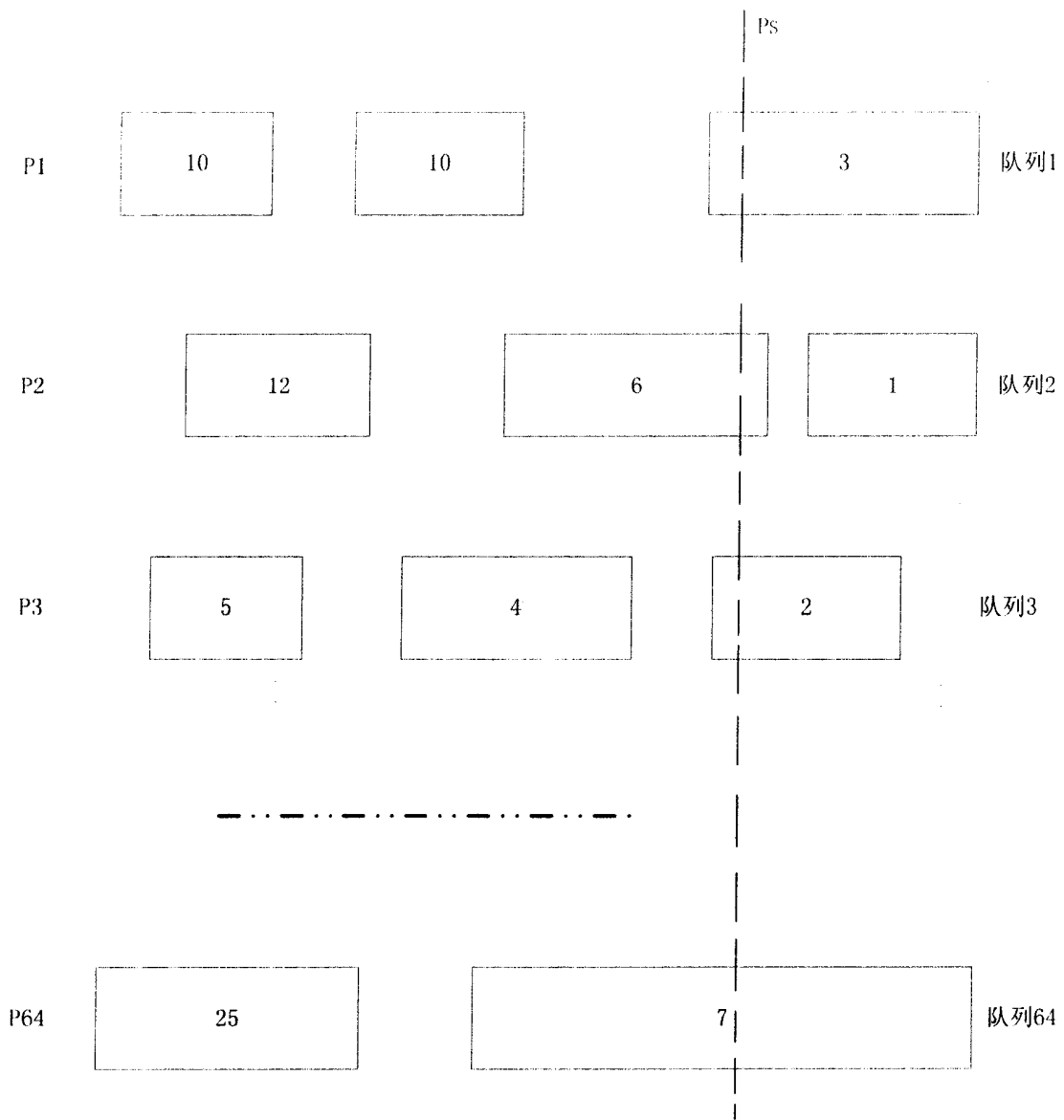


图2

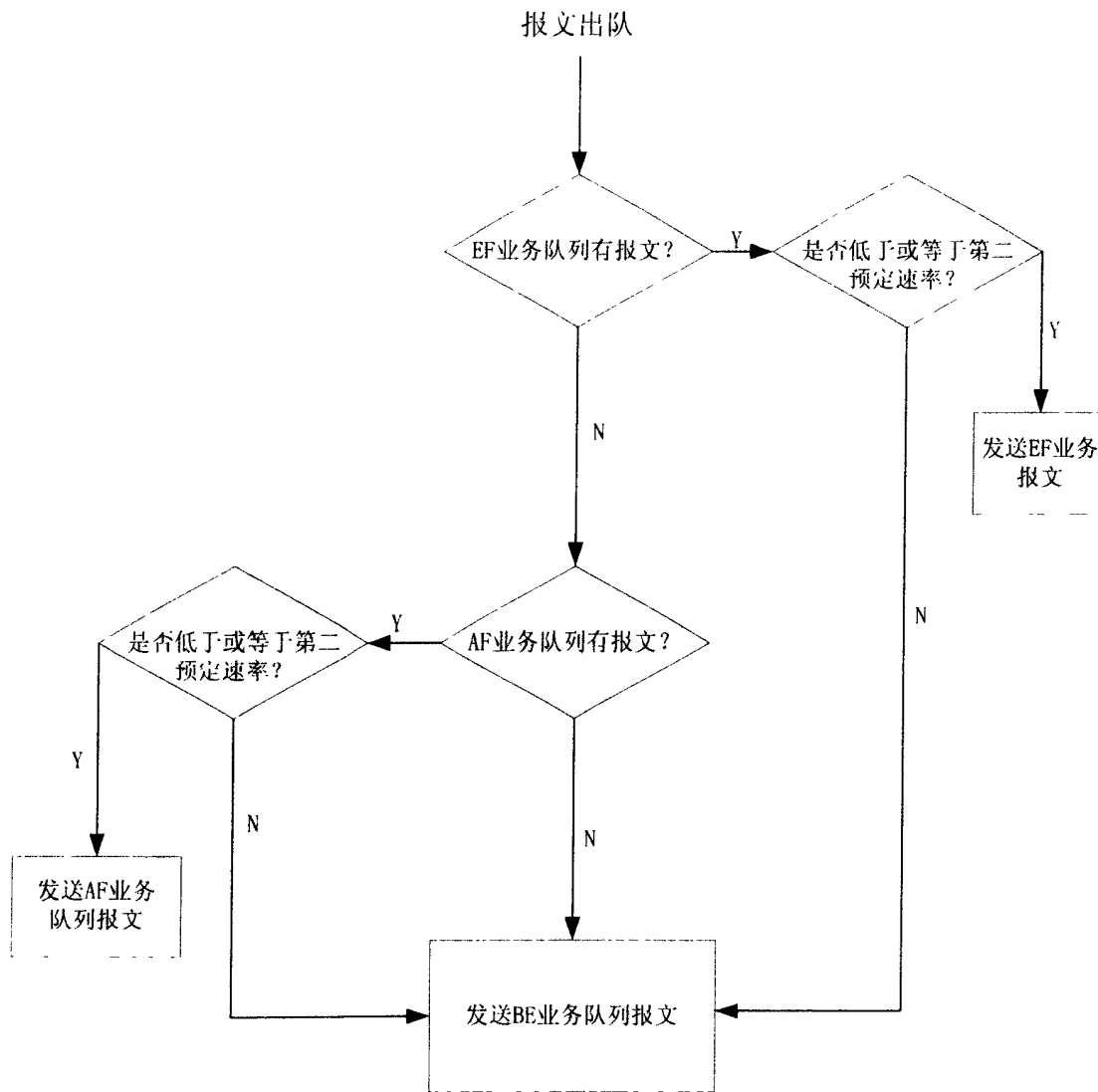


图3