

果这两个应用平等地共享该带宽,每个应用将会丢失其 25% 的传输分组。这是一个无法接受的低 QoS,以至于两个音频应用完全不可用;甚至从一开始就没有必要传输任何音频分组。

在图 7-27 中两个应用不能同时满足的情况下,这个网络应当做些什么呢?如果允许两者以一种不可用的 QoS 继续应用流上浪费网络资源,最终将无法向端用户提供任何效用。希望该答案是清晰的,即应用流中的一个应当被阻塞(即拒绝接入网络),而另一个应当允许继续进行,使该应用使用所需的全部 1Mbps。电话网是一个执行这种呼叫阻塞的网络例子,即如果要求的资源(在电话网的情况下是一个端到端的电路)不能分配给该呼叫,该呼叫就被阻塞了(阻止进入网络),并且返回给用户一个忙信号。在我们的例子中,如果一个流没有分配到足够的 QoS 来使自己可用,允许它进入网络没有任何好处。的确,接纳一个不能得到它需要的 QoS 的流是要付出代价的,因为网络资源被用于支持一个对端用户没有用的流。

通过基于流的资源要求以及已经准入流的资源要求明确地准入或阻塞流,网络能够保证准入流可得到它们所请求的 QoS。这隐含的意思是:需要为流提供有保证的 QoS 就需要流来申告它的 QoS 需求。让流申告它的 QoS 需求,然后让网络接受该流(以所要求的 QoS)或者阻塞该流的过程称为**呼叫准入**(call admission)过程。这则是我们的第 4 个见解(除了 7.5.2 节中前面 3 个见解外),该见解增加了提供 QoS 所需的机制。

**见解 4:** 如果充分的资源不总是可用的,并且要确保 QoS,则需要一个呼叫准入过程,在这个过程中流申告其 QoS 需求,然后要么被网络准入(以所要求的 QoS),要么被网络阻塞(如果网络不能提供所要求的 QoS)。

如果一个呼叫(一个端到端流)一旦开始就要被确保给定的服务质量,则我们在图 7-27 中的启发性例子强调了对几种新的网络机制和协议的需求:

- **资源预留。**为了满足一个呼叫所希望的 QoS,确保它具有所需的资源(链路带宽、缓存)的唯一方法是显式地为该呼叫分配这些资源,用网络的行话来讲,该过程称为**资源预留**(resource reservation)。一旦呼叫预约了资源,它在其整个过程中按需访问这些资源,而不管所有其他呼叫的需求。如果一个呼叫预约并得到链路带宽的  $x$ Mbps 的保证,而且传输速率决不大于  $x$ Mbps,则该呼叫将具有无丢包和无线延的性能。
- **呼叫准入。**如果预留了资源,则该网络必须具有一种用于呼叫请求和预留资源的机制。由于资源不是无限的,如果请求的资源不可用,则进行呼叫准入请求的呼叫将被拒绝准入,即被阻塞。电话网执行的是这种呼叫准入机制,即当我们拨一个号码时,就请求了资源。如果完成该呼叫所需的电路(TDMA 时隙)是可用的,则分配电路并且完成了呼叫。如果电路不可用,则该呼叫被阻塞,我们得到了忙信号。为了得到网络的准入,被阻塞的呼叫能够一再尝试,但是直到它成功地完成呼叫准入过程,才被允许向网络发送流量。当然,分配链路带宽的路由器,其分配的带宽不应当超过链路可用的带宽。通常,一个呼叫仅能预约一条链路的带宽的一部分,因此一台路由器可以为多于一条呼叫分配链路带宽。然而,如果要提供硬服务质量保证的话,为所有呼叫分配的带宽总和应当小于该链路的容量。