



[12] 发明专利说明书

[21] ZL 专利号 01805851.5

[45] 授权公告日 2004 年 9 月 22 日

[11] 授权公告号 CN 1168265C

[22] 申请日 2001.3.30 [21] 申请号 01805851.5

[30] 优先权

[32] 2000.3.31 [33] US [31] 09/540,428

[86] 国际申请 PCT/GB2001/001436 2001.3.30

[87] 国际公布 WO2001/076138 英 2001.10.11

[85] 进入国家阶段日期 2002.8.30

[71] 专利权人 国际商业机器公司

地址 美国纽约

[72] 发明人 布拉赫马南德·库马尔·高尔迪

黄东明 克拉克·德布斯·杰弗里斯

迈克尔·斯蒂芬·西格尔

卡迪克·苏迪普

审查员 冯于迎

[74] 专利代理机构 中国国际贸易促进委员会专利
商标事务所

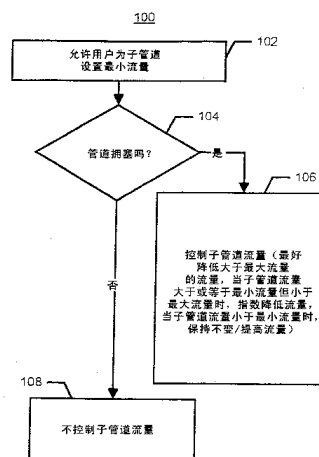
代理人 吴丽丽

权利要求书 5 页 说明书 19 页 附图 12 页

[54] 发明名称 在计算机网络的子管道中控制流量的
方法和系统

[57] 摘要

本发明公开了控制在计算机网络中的多个子管道的方法和系统。计算机网络至少包括一个交换机。多个子管道利用交换机通过网络传输通信量。该方法和系统包括允许为多个子管道的每一个设置最小流量和确定管道是否存在拥塞。该方法和系统还包括仅当拥塞存在时,控制多个子管道的子管道中流量。一旦确定拥塞存在,设置用于控制子管道的时间将被设置得充分长于为控制子管道中流量的单独降低或提高所需的时间。控制该流量,使得如果多个子管道的子管道流量小于最小流量,则子管道的流量可被线性提高。该流量控制也可使如果流量大于最小流量,则多个子管道的子管道流量被指数性降低。在决定是否提高或降低子管道中流量时,可使用拥塞的第二个定义。使得通过交换机的通信量稳定。



1. 一种方法，用于控制在包括至少一个交换机的计算机网络的管道中的多个子管道，多个子管道利用交换机通过网络传输通信量，该方法包括以下步骤：

- (a) 允许为多个子管道的每一个设置最小流量 (102)；
- (b) 确定管道是否存在拥塞 (104, 116)；
- (c) 如果拥塞存在，控制多个子管道中一个子管道的流量 (106, 132)，包括以下步骤：

(c1) 响应子管道流量小于最小流量的情况，使子管道的流量线性提高 (106) 或保持不变；和

(c2) 响应子管道流量大于最小流量的情况，使子管道的流量指数性降低 (106) 子管道流量。

2. 根据权利要求 1 所述的方法，还包括如下步骤：

- (b1) 确定是否存在额外带宽 (120)；

并且，

线性提高步骤 (c1) 还响应于额外带宽存在的情况；和

指数性降低步骤 (c2) 还响应于额外带宽不存在的情况。

3. 根据权利要求 2 所述的方法，其特征在于：额外带宽确定步骤 (b1) 还包括以下步骤：

(b1i) 将额外带宽值 (128) 设置为以前额外带宽值的指数性加权平均值。

4. 根据权利要求 2 所述的方法，其特征在于：交换机还包括具有队列的处理器，多个子管道使用队列通过交换机传输通信量，其中带宽确定步骤 (b1) 基于所确定的队列级别来确定额外带宽是否存在。

5. 根据权利要求 4 所述的方法，其特征在于：额外带宽确定步骤 (b1) 还包括以下步骤：

(b1i) 将额外带宽值 (128) 设置为以前额外带宽值的指数性加权平均值。

6. 根据权利要求2所述的方法，其特征在于：额外带宽确定步骤（b1）基于确定第二种类型拥塞是否存在来确定额外带宽是否存在，其中所述第二种类型拥塞基于管道拥塞级别（PCL）信号之外的参数来触发子管道的流量控制。

7. 根据权利要求6所述的方法，其特征在于：额外带宽确定步骤（b1）还包括以下步骤：

（b1i）将额外带宽值（128）设置为以前额外带宽值的指数性加权平均值。

8. 根据权利要求7所述的方法，其特征在于：多个子管道中子管道的流量是提供的速率乘以传输片断，且其中线性提高步骤（c1）将子管道传输片断（130）设置为以前的传输片断加上第一个参量，第一个参量是第一个常量乘以额外带宽值，和指数性降低步骤（c2）将子管道传输片断设置为以前的传输片断减去第二个参量，第二个参量是以前的流量乘以第二个常量。

9. 根据权利要求8所述的方法，其特征在于：第一个常量和第二个常量依赖管道的最小流量。

10. 根据权利要求8所述的方法，其特征在于：第一个常量是权重乘以第三个参量，第三个参量是队列服务速率加上管道最小流量减去多个子管道中每个的最小流量的总和。

11. 根据权利要求1所述的方法还包括以下步骤：

（a1）允许为多个子管道的每一个设置最大流量；

并且，控制步骤（c）还包括以下步骤：

（c3）如果子管道流量大于最大流量，则降低（106、130）子管道流量。

12. 根据权利要求1所述的方法，其特征在于：为多个子管道的每一个执行控制步骤（c）。

13. 根据权利要求1所述的方法，其特征在于：网络还包括边缘，且其中交换机位于网络边缘。

14. 根据权利要求1所述的方法还包括以下步骤：

(d) 在预先确定的时间周期后重复拥塞确定步骤(b)和在整个预先确定的时间周期中重复控制步骤(c)。

15. 前面任一个权利要求所述的方法, 其特征在于: 通过交换机的通信量包括多个分组, 并且其中拥塞确定步骤(b)基于用于多个分组中每一个的第一部分的显式拥塞通知(ECN)(140)的级别是否大于阈值来确定拥塞是否存在。

16. 根据权利要求15所述的方法, 其特征在于: 通过从子管道随机丢弃分组, 来降低多个管道的子管道中流量。

17. 根据权利要求15所述的方法, 其特征在于: 通过基于分组的优先级丢弃分组, 来降低子管道中的流量。

18. 根据权利要求1到14中任一个所述的方法, 其特征在于: 通过交换机的通信量包括多个分组, 并且其中拥塞确定步骤(b)基于多个分组(150)的一部分的级别是否大于阈值来确定拥塞是否存在, 多个分组的该部分包括多个同步分组。

19. 根据权利要求18所述的方法, 其特征在于: 通过从子管道随机丢弃分组, 来降低多个管道的子管道中流量。

20. 根据权利要求18所述的方法, 其特征在于: 通过基于分组的优先级丢弃分组, 来降低子管道中的流量。

21. 根据权利要求1到14中任一个所述的方法, 其特征在于: 通过交换机的通信量包括多个分组, 多个分组中的每一个由发送者发出并由接收者接收, 并且其中拥塞确定步骤(b)基于多个分组中每一个在发送者和接收者之间的往返时间的统计度量是否大于阈值, 来确定拥塞是否存在(160)。

22. 根据权利要求21所述的方法, 其特征在于: 通过从子管道随机丢弃分组, 来降低多个管道的子管道中流量。

23. 根据权利要求21所述的方法, 其特征在于: 通过基于分组的优先级丢弃分组, 来降低子管道中的流量。

24. 根据权利要求1到14中任一个所述的方法, 其特征在于: 通过交换机的通信量包括多个分组, 多个分组中每一个的第一部分由发

送者发出并由接收者接收，并且其中拥塞确定步骤(b)基于用于多个分组的每一个的第二部分的显式拥塞通知(ECN)、多个分组中每一个的第一部分在发送者和接收者之间的往返时间和多个分组的第三部分来确定一个拥塞度量，其中多个分组的第三部分包括多个同步分组(170)，通过比较该拥塞度量是否大于阈值来确定拥塞是否存在。

25. 根据权利要求24所述的方法，其特征在于：通过从子管道随机丢弃分组，来降低多个管道的子管道中流量。

26. 根据权利要求24所述的方法，其特征在于：通过基于分组的优先级丢弃分组，来降低子管道中的流量。

27. 一种控制计算机网络中的管道的多个子管道的系统，该计算机网络包括由多个子管道交叠的交换机，该系统包含：

在通过交换机传输通信量时被多个子管道使用的队列；和

排队机构，与队列相连，使用由用户为多个子管道的每一个设置的最小流量，来控制通过交换机的通信量，该排队机构为管道确定拥塞是否存在，并仅当拥塞存在时控制多个子管道中一个子管道的流量，使得如果多个子管道中子管道的流量小于最小流量，则子管道流量可以被线性提高，如果多个子管道中子管道的流量大于最小流量时，则该流量可以被指数性降低，使得通过交换机的通信量稳定。

28. 根据权利要求27所述的系统，其特征在于：排队机构还确定额外带宽是否存在，并且基于额外带宽是否存在为多个子管道的每一个提高或降低流量。

29. 根据权利要求28所述的系统，其特征在于：排队机构通过确定第二种类型拥塞是否存在，来确定额外带宽是否存在，如果第二种类型拥塞存在，则指示额外带宽存在，否则，指示额外带宽不存在，其中所述第二种类型拥塞通过管道拥塞级别(PCL)信号以外的参数来触发子管道的流量控制。

30. 根据权利要求27所述的系统，其特征在于：通过交换机的通信量包括多个分组，并且其中排队机构基于用于多个分组中每一个的第一部分的显式拥塞通知(ECN)的级别是否大于阈值，来确定拥塞

是否存在。

31. 根据权利要求 27 所述的系统，其特征在于：通过交换机的通信量包括多个分组，并且其中排队机构基于多个分组的一部分的级别是否大于一个阈值来确定拥塞是否存在，多个分组的该部分包括多个同步分组。

32. 根据权利要求 27 所述的系统，其特征在于：通过交换机的通信量包括多个分组，多个分组的每一个由发送者发出并由接收者接收，并且其中排队机构基于多个分组的每一个在发送者和接收者之间的往返时间的统计度量是否大于一个阈值来确定拥塞是否存在。

33. 根据权利要求 27 所述的系统，其特征在于：通过交换机的通信量包括多个分组，多个分组中每一个的第一部分由发送者发出并由接收者接收，并且其中排队机构基于用于多个分组的每一个的第二部分的显式拥塞通知（ECN）、多个分组中每一个的第一部分在发送者和接收者之间的往返时间和多个分组的第三部分来确定一个拥塞度量，所述多个分组的第三部分包括多个同步分组，然后通过比较该拥塞度量是否大于一个阈值，确定是否存在拥塞。

34. 根据权利要求 27 所述的系统，其特征在于：交换机包括对应于多个刀形开关的多个处理器，多个处理器的每一个具有多个端口，并且其中队列用于多个处理器中的处理器。

35. 根据权利要求 27 所述的系统，其特征在于：网络包括边缘，并且交换机位于网络边缘。

在计算机网络的子管道中控制流量的方法和系统

技术领域

本发明涉及计算机网络，尤其涉及为计算机网络提供分类服务和较好控制的方法和系统。

背景技术

由于各种网络应用（如，那些涉及因特网的网络应用）的使用日益增长，计算机网络正在引起人们越来越大的兴趣。图 1 描述了通常的网络 10 和 30 通过因特网 1 相连。网络 10 和 30 分别包括主机 11、12、17、18、21、22 和 32、34。每个网络 10 和 20 还分别包括交换机 14、16、19 和 36，并且还可能分别包括一个或多个服务器，如服务器 15、20、24 和 38。另外，每个网络 10 和 30 还可能分别包括一个或多个网关 10 和 35，以连接因特网 1。图中未示出的路由器和网络 10 和 30 的其他部分，也可能通过网络 10 和 30 控制通信量，并且通常被认为自然由交换机 14、16、19 和 36，以及网络 10 和 30 描述。另外，因特网 1 包括它自己的交换机和网络，图中也未明确示出。

交换机，如交换机 14 或 36 和因特网 1 中与诸如交换机 14 或 16 相连的交换机（未示出）分别被认为在网络 10、30 或因特网 1 的边缘。这是因为这些交换机直接向实体发送通信量和/或直接从实体接收通信量，而不是直接在网络 10、30 或因特网 1 的控制下。例如，一个网络管理员或网络 10 的其他用户可以为网络 10 的使用和执行控制参数设置。然而，一个典型的网络管理员并不控制因特网 1，也不可以控制网络 30。交换机 14 直接与网关 13 相连，以提供对因特网 1 的访问。而交换机 16 和 19 却不是这样，因此交换机 13 被认为在网络 10 的边缘。类似地，因特网 1 的一个交换机（未示出）也

与网关 13 对接。一个网络管理员或其他用户可以控制因特网 1 的一些或全部运行，但却控制不了网络 10 的相应部分。这样一个交换机将被认为在因特网的边缘。

图 2 描述了能够在计算机网络中使用的交换机 40 的高级框图。这样，交换机 40 可以在因特网 1 中使用，和为交换机 14、16、19 和 36 使用。交换机 40 包括一个与刀形开关 47、48 和 49 相连的交换机构 41。每一个刀形开关 47、48 和 49 通常是一个电路板并包括至少一个与端口 44 相连的网络处理器 42。这样，端口 44 可与交换机所在的网络中的主机或其它部件相连。刀形开关 47、48 和 49 可提供进入交换机构 41 的通信量并接收来自交换机构 41 的通信量。因此，与刀形开关 47、48 或 49 中任一相连接的任何部件均可与连接在另一个刀形开关 47、48 或 49 的另一个部件通信，或与和自身相连的所述刀形开关相连接的另一部件通信。

图 3A 描述了交换机 40 的另一简化的框图，该图阐明了网络处理器 51 和 55 执行的一些功能。交换机 40 连接与端口 A52 连接的部件和与端口 B76 连接的部件。交换机 40 执行各种功能，包括进入交换机 40 的数据分组的分类，传输通过交换机 40 的数据分组和重装数据分组。这些功能分别由分类器 58、交换机构 64 和重装器 70 提供。分类器 58 将进入的分组分类并分割每个分组为适宜长度的片段，该片段称为信元。交换机构 64 是一个连接矩阵，信元被通过它按照各自的路径传输通过交换机 40。重装器 70 重装信元为正确的分组。分组可被提供给适当的端口 B76，并输出到目的主机。分类器 54 可以是网络处理器 51 的一部分，同时重装器 70 可以是另一个网络处理器 55 的一部分。上述的网络处理器 51 和 55 分别执行从端口 A52 输出并向端口 B76 输入通信量的功能。而且，网络处理器 51 和 55 也分别执行从端口 B76 输出通信量和输入到端口 A52 的功能。因此，每一个网络处理器 51 和 55 能执行分类和重装的功能。

由于在传输通信量通过交换机 40 时的瓶颈，可以要求数据分组先等待分类、传输和重装功能的执行。结果是，产生队列 56、62、

68 和 74。与队列 56、62、68 和 74 相连的是排队机构 54、60、66 和 72。排队机构 54、60、66 和 72 把分组或信元放置到相应的队列 56、62、68 和 74 中并提供一个通知发送回产生分组的主机。

尽管队列 56、62、68 和 74 被独立描述，但技术中的通常技巧将很容易认识到队列 56、62、68 和 74 中的一些或全部可以是同一物理存储器资源的一部分。图 3B 描述了这样一个交换机 40'。交换机 40' 的许多部件与交换机 40 的部件是类似的。这些部件也因此有类似标识。例如，在交换机 40' 中的端口 A52' 对应与交换机 40 中的端口 A52。在交换机 40' 中，队列 54 和队列 62 共享一个单独的存储器资源 59。类似地，队列 68 和队列 74 是另一个单独的存储器资源 71 的一部分。这样，在交换机 40' 中，队列 56、62、68 和 74 是从存储器资源 59 和 71 划分的逻辑队列。

目前，大多数常规交换机 40 处理通过网络的通信流量，在网络中交换机的使用是相同的。然而，有一种为客户提供不同服务的趋势，如基于客户为服务所支付的费用。一些客户可能希望多花费用以确保较快速的响应或确保即使因拥塞造成其他客户的通信量下降时他的通信量也能被传送。因此，已经开发了分类服务(differentiated services)的概念。分类服务能提供不同的服务级别，或为不同的客户提供不同的通过网络的通信流量。

DiffServ 是一个为提供分类服务而制定的因特网网络工程部分(IETF)标准(参见 IETF RFC 2475 和相关的 RFCs)。DiffServ 基于行为聚集流。一个行为聚集流可视为从一个网络边缘到另外一个网络边缘的管道。在每一个行为聚集流中，可能有单独主机间的数百个会话。然而，DiffServ 并不关心一个行为聚集流中的会话。相反，DiffServ 关心的是行为聚集流之间的带宽分配。根据 DiffServ，额外的带宽在行为聚集流之间公平地分配。而且，DiffServ 提供用于测量提供给每一个行为聚集流的服务级别的标准，这在下面讨论。

提供不同服务级别的一个通常机制是利用权重和队列级别的组合，以提供不同服务级别，例如，San Jose, California 的 Cisco Systems

Inc.提供的产品。图 4 描绘了这样一个通常的方法 80。队列级别阈值和权重通过步骤 82 设置。典型的情况是，在步骤 82 中通过网络管理员调整或“调节”来设置队列级别阈值。通过某一队列、交换机 40 或网络处理器 42 可以为不同的管道或流量设置权重。因此，权重典型的是为不同的行为聚集流设置。通过步骤 84，典型的情况下在一个时间周期的末端观察队列级别。通过步骤 86，然后，基于队列级别与队列级别阈值的比较结果和权重改变管道流。具有较高权重的管道流在步骤 86 中经历较大的变化。管道流确定提供给一个队列，如队列 56 的通信量，通过相应的排队机构，如排队机构 54 通过管道被传送到队列 56 中。通过步骤 88，通信量就这样基于流量被传送到队列或被丢弃。通过步骤 90，网络管理员然后确定当前的服务级别是否理想。如果理想，网络管理员就完成了拥塞避免。然而，如果没有达到理想的服务级别，则通过步骤 82 重新设置队列级别阈值、或者权重，并重复方法 80。与常规方法 80 相关的信息可以在 <http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cqcr/qosc/qcpart2/qcconman.htm> 中找到。

尽管常规方法 80 能发挥作用，但技术中的通常技巧将很容易认识到通过网络确定改变队列级别阈值将对某管道的影响是困难的。因此，网络管理员使用方法 80 在为不同的客户或管道（行为聚集流）在一个计算机网络中达到理想的流量速率前，不得不做大量的试验。

而且，方法 80 间接操作典型用于测量服务质量的参数。队列级别不是一个典型的用于服务的测量标准。典型的是，例如在 DiffServ 中（参见 IETF RFC 2475 和相关的 RFCs），服务级别通过四个参数测量：丢弃率、带宽、等待时间和抖动。丢弃率是指在通信量通过交换机时被丢弃的百分比。一个行为聚集流的带宽是对通过交换机并到达目的地的行为聚集流的通信总量的测量。等待时间是通过网络发送通信量中出现的延迟。抖动是等待时间随时间的改变。队列级别不被认为是一个服务质量的直接测量。因此，方法 80 并不直接提出针对服务质量的任何标准。因此，对网络管理员来说，利用方法 80 提供不同的

服务级别也更加困难。另外，方法 80 为行为聚集流进行流量控制。方法 80 并不能在一个比这更好的级别上控制通信量。

其他常规的控制通信量的方法为利用流量、最小流量速率、权重、优先级、阈值和一个指示额外带宽或传输通信量能力存在的信号来控制流量。然而，常规的方法是基于直觉，且表面看来有说服力，但没有实际上的基于控制理论的依据支持。这种常规方法控制通信量通过交换机是否是一个稳定的机制还不清楚。因此，这种常规方法不能够充分控制通信量通过交换机 40。

另外，另一个 IETF 建议是适用集成服务的 IntServ。在 IntServ 中，在任何时候网络中每个流量总是被控制。因此，IntServ 提出控制网络的每一个行为聚集流或管道中的每一个单独的流量。结果是，IntServ 可被用来提供分类服务。然而，快速地在全部时间控制所有的流量是不切实际的。尤其是，随着网络和通过网络的通信量的增长，在网络中流量数也快速增长。对于大多数的网络，在全部时间控制网络中的每一个流量将耗费巨大的资源，且很难完成。因此，IntServ 也不能充分控制通过网络的通信量。

因此，所需的是为提供更好的分类服务和在一个较好水平上控制流量的系统和方法。本方法专注于这样一种需要。

发明内容

本发明提供了控制在计算机网络中的多个子管道的方法和系统。计算机网络包括至少一个交换机。多个子管道利用交换机通过网络传输通信量。该方法和系统包括允许为多个子管道的每一个设置最小流量和确定管道是否存在拥塞。该方法和系统还包括只有当存在拥塞时控制多个子管道的子管道中的流量控制。该流量控制在多个子管道的子管道流量小于最小流量的情况下，子管道的流量可线性提高。该流量控制在流量大于最小流量的情况下，多个子管道的子管道流量可指数性降低。使得通过交换机的通信量稳定。

按照这里公开的系统和方法，当考虑分类服务时，本发明以一个

较好水平提供一种稳定、可行的机制以控制通过一个网络的通信量。另外，本发明可被用来在网络边缘控制通过网络的通信量，因此，允许在网络内部消除冗余控制。

附图说明

本发明的优选实施例将参照下图描述：

图 1 是根据现有技术的计算机网络的架构图；

图 2 是根据现有技术的交换机的高级结构图；

图 3A 是根据现有技术的交换机的简化结构图；

图 3B 是根据现有技术的另一个交换机的简化结构图；

图 4 是描述根据现有技术通过交换机提供不同级别服务的常规方法的流程图；

图 5 是描述本发明可被使用的网络图；

图 6 是描述为控制通信量并在子管道中提供不同级别服务，依照本发明方法的高级流程图；

图 7A 和 7B 描述了为控制通信量并在子管道中提供不同级别服务，依照本发明方法的更为详细的流程图；

图 8 是描述为确定管道中是否存在拥塞，依照本发明方法的首选实施例的详细流程图；

图 9 是描述为确定管道中是否存在拥塞，依照本发明方法的第二个实施例的详细流程图；

图 10 是描述为确定管道中是否存在拥塞，依照本发明方法的第三个实施例的详细流程图；和

图 11 是描述为确定管道中是否存在拥塞，依照本发明方法的第四个实施例的详细流程图；

具体实施方式

本发明将根据特定系统和特定部件加以描述。然而，本领域技术人员将很容易认识到该方法和系统针对计算机网络中的其他部件也能

有效运作。本发明还将在队列的上下文里描述。然而，本领域技术人员将很容易认识到当队列是单个存储器资源的逻辑队列部分时或当队列是分离存储器资源的一部分时，本发明也可有效运作。而且，当控制流量进入一个特定逻辑队列的子队列时，本发明可类似运作。本发明也将根据通过主动地丢弃分组来控制网络流量加以讨论。然而，本领域技术人员将很容易认识到该方法和系统控制到达队列中的分组的速率，并且发送到源并指示将要发送的分组片断的信号连同源中的其它信号将是有效的。因此，传输片断的控制类似于例如由一个或多个源提供分组的速率的控制。而且，本发明将针对到一个队列的管道和子管道加以描述。然而，本领域技术人员将很容易认识到管道可能是流入利用队列进行存储的特定部件中的不同或相同种类、或任何提供速率的行为聚集流。本发明也将在控制特定管道的子管道的上下文里加以描述。然而，本领域技术人员将很容易认识到本发明也可用于在多个管道的子管道里控制流量。

现在参照图 5 可获得依照本发明的方法和系统的更详细的阐述，该图描绘了本发明可被使用的优选实施例网络 10' 和 30' 以及因特网 1。网络 10' 和 30' 以及因特网 1' 基本上与图 1 中的网络 10 和 30 以及因特网 1 相同。然而，也示出了控制点 25、26、27 和 37。控制点 25、26、27 和 37 最好是与交换机 14'、16'、19' 和 36' 连接的普通用途计算机。控制点 25、26、27 和 37 最好执行与矩阵和在检测过滤规则中使用的位选择相关的功能，这在下面讨论。交换机 14'、16'、19' 和 36' 最好包括软件管理决策树（图 5 中未示出），这在下面讨论，并使用决策树确定一个关键字是否匹配一个或多个过滤规则。另外，交换机 14'、16'、19' 和 36' 最好与图 2、3A 和 3B 中描述的交换机 40 和 40' 一样。尽管网络 10'、30' 和因特网 1' 优先使用，但网络 10 和 30 以及因特网 1 也可结合本发明使用。

现在参照图 6 可获得依照本发明的方法和系统的更详细的阐述，该图描绘了依照本发明方法 100 的实施例。方法 100 最好使用美国专利申请号 No.09/384, 691, 1999 年 8 月 27 日申请，标题为

“ NETWORK PROCESSING COMPLEX AND METHODS” 中公开的设备完成，该申请转让给本申请的受让人。

可通过图 2、3A 和 3B 中所示的交换机 40 或 40'使用方法 100。因此，方法 100 可在具有多刀形开关 47、48、49 和在每个刀形开关具有多个端口的交换机中执行。为清楚起见，方法 100 将结合图 3A 中描述的队列 74 和排队机构 72 加以解释。然而，可通过其它队列，如队列 56、62、68、56'、62'、68'和 74'，使用方法 100。也可通过其它排队机构，如排队机构 54、60、66、54'、60'、66'和 72'，使用方法 100。在优选实施例中，方法 100 在多队列是相同存储器资源一部分的系统中使用。然而，方法 100 也可在每个队列具有独立存储器资源的另一个系统中使用。方法 100 最好在位于网络边缘的交换机中使用，如图 1 和 5 中描述的交换机 14 或 14'。另外，方法 100 最好仅使用于位于网络边缘的交换机中，而不在网络内部的交换机中使用。另外，尽管方法 100 最好用于控制到达队列的子管道并使用排队机构执行，但并不阻止该方法被交换机的其它部分使用。

方法 100 用于在网络中为子管道控制流量。每个子管道是管道的一部分。管道也通过网络流动。因此，特定网络的管道可以被认为是网络边缘终止。最好，每个管道都是一个行为聚集流。一个子管道是管道中多个流量的组合。因此，子管道也可以被认为是网络边缘终止。换句话说，管道和子管道定义在网络内。一个子管道可能包括一个单一流量或包括管道中多个流量的组合。方法 100 还最好仅在网络边缘的交换机中使用，在所述网络边缘处子管道和管道终止。然而，方法 100 也可使用在网络内部的交换机中。

参照图 3A 和 6，通过步骤 102，可设置提供通信量到交换机 40 的特定管道的子管道的最小流量。在步骤 102 中设置的最小流量最好是保证提供给予管道的最小流量。因此，在步骤 102 中设置的最小流量对应于子管道总是被允许消耗的带宽总量。因此，最小流量最好这样设置，如果每个子管道传送的流量等于最小流量时，管道将不拥塞并且通信量将以理想的参数传播通过交换机和网络。当通过子管道的

通信量比较小时，子管道可传送小于最小流量的通信量。子管道也可以被允许传送比最小流量多的通信量。也是在优选实施例中，可以设置最大可能流量，可能用于拒绝服务攻击。在优选实施例中，为进入某一队列的子管道设置最小流量，如队列 74。另外，每个子管道最好以某一给定速率， $I_i(t)$ ，向队列 74 提供通信量，这里， i 代表第 i 个子管道。依赖不同的因素，一些通信量可能被丢弃而不是进入队列。第 i 个子管道的传输片断， $T_i(t)$ ，是来自传送到队列 74' 的第 i 个子管道的通信量的片断。因此，第 i 个子管道的瞬时流量是 $f_i(t) = I_i(t) * T_i(t)$ 。

通过步骤 104，确定其子管道受控的管道是否拥塞。某一管道可基于多种因素被确定为拥塞。典型地基于流过管道的分组或通信量来估计这些因素。注意，这里使用的分组是指各种类型的数据分组，包括但不限于以太网分组（通常称帧），ATM 分组（通常称信元）和 IP 分组（通常称分组）。确定管道中是否拥塞的一些因素包括某一分组从发送者到达接收者以及接收者发回确认所花费的时间、同步分组的数目、往返时间（RTT）和显式拥塞通知（ECN），这些因素可以为分组提供，下面参照图 8-11 讨论。返回参照图 6，一般说来，当管道传送足够大的通信量，以致于以一个如网络管理员的用户认为不能接受的比率延迟或丢弃分组时，管道可以被认为拥塞。

如果确定拥塞不存在，通过步骤 108，子管道中的流量不被控制。因此，当管道的执行被认为可以接受时，交换机 40 并不需要扩充资源以控制子管道中的流量。然而，如果确定拥塞存在，那么通过步骤 106，子管道中的流量被控制特定的时间， T 。子管道控制的特定时间 T 典型是用于更新传输片断的时间增量 Dt 的许多倍（也许是一千或一百万倍）。这样，在 T 期间，如果拥塞信号当前指示拥塞，那么，只要拥塞持续且流量保持在保证最小速率之上，那些在其保证最小速率之上的子管道中的流量就呈指数性减少。另外，注意 T 和 T_i 的区别， T 是指子管道中控制流量的时间间隔， T_i 是指为某一子管道的传输片断。在优选实施例中，步骤 106 包括指数性降低流量大于相应于在步骤 102

中设置的最小流量的子管道中的流量。因此，子管道的流量与子管道流量值成比例降低。也在优选实施例中，步骤 106 可包括使其流量小于或等于相应最小流量的子管道流量保持不变，至少是在开始时。然而，在时间间隔 T 期间，如果拥塞不被即刻指示，那么那些在其最大速率之下的子管道的流量被允许线性提高。特别地，当子管道继续被控制时，那些其流量在设定的最小值之下的子管道保持不变或允许线性提高，那些如果在其最大流量（如果设置的话）之上的子管道将被迫降低，和那些其流量位于最小和最大值之间的子管道被允许线性提高或被迫指数性降低。在步骤 106 继续控制流量时，尤为如此。因为流量被指数性降低或被线性提高，子管道的流量以一个稳定的方式被控制。步骤 106 最好包括为被控的子管道设置传输片断。也在步骤 106 中，使用传输片断控制进入队列子管道的分组数目。

因此，方法 100 仅在管道拥塞时控制子管道中的流量。结果是，方法 100 在拥塞不存在时不会不必要地消耗资源，因此，通信量理想地流过网络。而且，因为方法 100 可线性提高和指数性降低子管道中的流量，通过子管道的流量控制将是稳定的。方法 100 也最好仅在网络边缘执行，比如在图 5 中描绘的网络 10' 的交换机 14' 中。然而，管道和子管道因此流过整个网络 10'。因此，在网络边缘，图 6 的方法 100 通过控制子管道控制网络的整个管道的拥塞。方法 100 不必在网络中的其它点执行。因此，在计算上昂贵和不必要的子管道冗余控制可以被避免。然而，即使当其它机构在网络中心和/或边缘执行时，方法 100 也可结合这种其它机构使用，以在网络中控制通信量。

图 7A 和 7B 描绘了根据本发明用于在子管道中控制流量的方法 110 的一个实施例更为详细的流程图。方法 110 最好在为被控的子管道设置最小流量后开始。

方法 110 将在为某一管道的子管道控制流量的上下文里加以描述。然而，本领域技术人员将很容易认识到方法 110 可扩充为与多管道一同使用。另外，方法 110 类似于方法 100，并因此可使用相同的设备执行。设置子管道的最小流量后，通过步骤 112 基于子管道的最

小和最大（如果设置）流量为每个子管道计算常量。对每一个子管道 i ，在步骤 112 中计算常量 C_i 和常量 D_i 。常量 C_i 用于为子管道 i 线性提高流量，这在下面讨论。类似地，常量 D_i 用于为子管道 i 指数性降低流量，这在下面讨论。在优选实施例中，常量 C_i 和 D_i 都是基于最小流量。在另一个实施例中，可以为不同的子管道提供权重。在这种情况下，常量 C_i 和 D_i 也可基于提供的权重计算。

一旦常量确定，通过步骤 114，通信量允许通过子管道流向队列 74。通过步骤 116，然后确定管道是否拥塞。在优选实施例中，步骤 116 通过使用每个被监视管道的管道拥塞级别（PCL）信号执行。管道的 PCL 位于 0 和 1 之间。PCL 越接近 1，管道越拥塞。因此，0 和 1 之间的阈值可被设置为拥塞阈值。步骤 116 将确定管道的 PCL。如果管道的 PCL 小于或等于这个阈值，步骤 116 将指示拥塞不存在。如果 PCL 大于这个阈值，那么步骤 116 将确定拥塞在管道中存在。如果没有拥塞，那么，通过步骤 134 确定时间段是否结束，在该时间段中可能希望执行控制。因此，如果没有拥塞，那么子管道流量可以不被控制。

如果在步骤 116 中确定拥塞存在，那么通过步骤 118 和 132 控制子管道中的流量。如果不是已经可用，通过步骤 118，为以前的时间段确定队列级别，一个瞬时额外带宽信号 B ，用二进制值 0 或 1 表示，和额外带宽值 E ，用 0 和 1 之间的数值表示。在优选实施例中，在步骤 118 也使以前的时间段的子管道流量可用。换句话说，在步骤 118 确定在方法 110 中更新系统所必须的数量。在优选实施例中，步骤 118 通常是不必要的，因为上面提及的数量在以前的时间段使用方法 110 确定。队列级别是子管道正在流向其中的队列的级别。队列级别可定义为最大队列级别的片断。例如，队列级别可以是图 3A 的队列 74 的满级别的一部分。返回参照图 7A 和 7B，在优选实施例中，被确定的队列级别是整个存储器资源的级别。然而，阻止不了为逻辑队列或子队列确定队列级别。在下面讨论瞬时额外带宽信号， B ，用二进制值 0 或 1 表示，和额外带宽值， E ，用 0 和 1 之间的数值表示。

通过步骤 120 确定是否存在额外带宽。在一个实施例中，只有当

如队列 74 的队列级别是 0 或低于一个小值或正在降低时,确定额外带宽存在。然而,确定是否存在额外带宽最好依赖于往返时间或其它参数、或在这种测量中的变化的单一确定。如下描述,这些其它参数也可用于测量拥塞。因此,在优选实施例,可认为使用了拥塞的两个定义。拥塞的第一个定义,在步骤 116 中,被用来触发子管道的流量控制。拥塞的第二个定义可用在步骤 120 中,以确定额外带宽是否存在。在步骤 120 中使用的拥塞的第二个定义可以不同于拥塞的第一个定义。例如,在步骤 120 中拥塞的第二个定义,在流量速率低于步骤 116 中使用的第一个拥塞定义时,可认为拥塞存在。另外,额外带宽的定义,和拥塞的第二个定义,可随时间间隔 T 而改变。例如,在时间间隔 T 的早期,额外带宽的定义可能包括拥塞测量,以确保子管道中的流量将被控制以减少拥塞。在时间间隔 T 的晚期,额外带宽定义可能改变以排除拥塞,为被认为存在的拥塞提高所需流量和/或仅依赖于其它参数,比如队列级别。特别地,队列级别和最大缓冲区容量(称为净空间)的区别可用于定义拥塞。然而,阻止不了用于确定额外带宽是否存在的另一个标准的使用。一般说来,瞬时额外带宽信号指示在以前的时间段额外资源是否存在于交换机 40,或整个管道中,以处理额外的通信量。如果额外带宽不存在,那么瞬时额外带宽信号, B , 通过步骤 122 设置为 0。信号 B 被称为瞬时的,是因为它基于队列级别、往返时间、或其它参数的单一测量或在这种测量中的变化的单一确定。这样,因为额外带宽的定义可依赖于拥塞,所以,瞬时额外带宽信号, B , 也可依赖于用于确定额外带宽是否存在的拥塞。通过步骤 124, 额外带宽值, E , 指数性提高到 1 (例如,如果没有拥塞)或降低到 0 (例如,如果有拥塞)。在优选实施例中,额外带宽值是瞬时额外带宽信号的指数性加权平均值。因此,额外带宽值提供在以前的时间段上可获得的资源的度量。另外,因为额外带宽值, E , 依赖于瞬时额外带宽信号, B , 所以,额外带宽值, E , 也可依赖于拥塞的存在,如用于确定是否存在额外带宽那样。

如果在步骤 120 中确定额外带宽存在,那么,通过步骤 126,瞬

时额外带宽值被设置为 1。通过步骤 128，额外带宽值被设置为瞬时额外带宽信号的指数性加权平均值。在优选实施例中，额外带宽值是第一个常量乘以以前的额外带宽值，再加上第二个常量乘以瞬时额外带宽值。在步骤 122 和 124 或步骤 126 和 128 中设置的瞬时额外带宽信号，B，和额外带宽值，E，的值将最好使用于下一个时间段，以为子管道更新流量控制。

在优选实施例中，通过步骤 130，与步骤 120 到 128 并行地为每个子管道 i 设置传输片断。然而，在另一个实施例中，可以与瞬时额外带宽信号，B，和额外带宽值，E 串行地更新传输片断。如果管道流量在最大级别以上，那么，最好通过设置传输片断为 $31/32$ 乘以以前的传输片断来指数性降低流量。如果管道流量不在最大级别以上，那么，传输片断如下改变。如果以前的瞬时额外带宽信号，B，是 1（额外带宽可获得），那么，步骤 130 将基于以前的传输片断、常量 C_i 、和额外带宽值为每一个子管道 i 设置传输片断。最好，在步骤 130 中为以前存在的额外带宽设置的传输片断是：

$$T_i(t+Dt) = T_i(t) + C_i * E(t)Dt$$

其中：

Dt = 时间段的长度（自传输片断上次被计算以后的时间）

最好，设置流量单元以使 Dt 和最大可能队列级别， Q_{max} ，为 1。因此，如果额外带宽存在，子管道的传输片断可线性提高。而且，只要额外带宽继续存在，传输片断就可以继续线性提高。然而，注意在优选实施例中，当子管道首次被控制时（当第一次确定管道中拥塞存在），通常将没有额外带宽。因此，子管道流量将不提高。因此，在步骤 116 中做出的控制子管道的决定是基于管道中拥塞开始的一个定义。子管道控制实施时间间隔 T，该时间间隔 T 通常是时间段的时间增量 Dt 的许多倍。最初，在 T 期间，子管道流量很可能被迫降低，以响应上面通常基于管道拥塞第二个定义的机制。拥塞第二个定义，至少在时间间隔 T 的早期用于确定是否存在额外带宽。然而，在时间间隔 T 中，可能存在子管道流量速率实际提高的时刻。在经过一些时

间段后, 接近时间间隔 T 末端时尤其如此。因此, 拥塞的一个定义可用于开始子管道中的流量控制。拥塞的第二个定义在定义额外带宽中被使用, 以允许瞬时额外带宽信号, B , 是 0 或 1, 从而允许子管道中的流量在时间间隔 T 期间即可提高又可降低。

如果在以前的时间段 (B 是 0) 额外带宽不存在并且子管道传送比它最小流量大的通信量, 那么, 在步骤 130 中子管道 i 的传输片断 $T_i(t+Dt)$ 基于子管道以前的传输片断 $T_i(t)$ 、 D_i , 和以前的子管道流量 $f_i(t)$, 被设置。若 $B = 0$ 且子管道传送比它最小流量大的流量, 则传输片断最好通过下式获得:

$$T_i(t + Dt) = T_i(t) - D_i * f_i(t)$$

另外, 当 $B = 0$ 时, 若子管道传送它的最小流量或更少的流量, 则传输片断最好通过下式获得:

$$T_i(t + Dt) = T_i(t)$$

因此, 在步骤 130 中设置的传输片断确保, 只要额外带宽继续不存在, 传送大于它们最小流量的子管道的传输片断, 和这些子管道的流量就呈指数性降低。常量 C_i 和 D_i 最好基于最小流量值。另外, 在优选实施例中, 当控制第一次施加时, 由于管道中的拥塞, 额外带宽将不存在。因此, 超过它们最小流量的子管道的流量将通过使用传输片断被指数性降低。没有超过它们最小流量的子管道的流量将在额外带宽不存在时保持不变。由于子管道最小值的设置最好是, 如果每个子管道传送等于最小流量的流量, 拥塞将不存在, 所以拥塞因此得到控制。从而, 若拥塞存在, 至少有一个子管道传送比它最小流量大的流量。因此, 指数性减少这些子管道的流量可纠正拥塞。如上讨论, 当额外带宽可得到时, 所有子管道的流量可线性提高。因此, 拥塞可被控制且子管道的流量以一种稳定的方式调节。

使用在步骤 130 中计算的传输片断, 通过步骤 132, 经过交换机 40 的分组在时间段期间被传输或丢弃。分组最好通过不允许分组进入某一队列, 如队列 74, 被丢弃。在优选实施例中, 分组的丢弃不仅基于其流经管道的传输片断, 而且基于每个分组的优先级。在另一个实

施例中，分组被随机丢弃。另外，可使用丢弃片断，而不使用传输片断。子管道丢弃片断是 1 减去传输片断。这样，丢弃片断指示，为使某一子管道获得理想的传输片断，或通信流量，应该被丢弃的分组片断。

通过步骤 134，确定时间段是否完成。如果没有完成，通过步骤 132，基于同样的传输片断继续传输通信量。如果时间段已经完成，那么，重复从步骤 116 开始的方法。然而，在一个实施例中，当方法 110 重复时，步骤 116（确定拥塞是否存在）可跳过某一时间周期或多个时间段。跳过步骤 116 允许子管道中的流量控制在拥塞不再存在之后，继续某一时间量。换句话说，跳过步骤 116 允许被设置用来控制子管道中流量的时间， T ，截止。在这样一个实施例中，允许高度控制子管道，并且允许提高子管道中的流量以确保额外带宽的分配。

方法 110 执行的时间段长度， Dt ，最好在方法 110 开始之前被设置。然而，在另一个实施例中，时间段的长度可能改变。另外，时间段的时间标度最好相对长些，或许是 1 到 10 毫秒或更长。应该设置时间段长度，以说明管道和子管道末端之间的延迟。这是因为在方法 100 中子管道的控制，根据是否存在拥塞和可否得到额外带宽，在时间段末端可以改变。为确保在一个时间段执行的控制有机会清除管道中的拥塞和能够给方法 100 或 110 提供正确的反馈，分组应该有充足的时间到达目的地并提供关于是否已经到达目的地的通知（如果有的话）。换句话说，时间段的长度应该足够长，以允许传输片断中的改变生效并向系统提供关于传输片断的影响的信息。

因为方法 110 线性提高流量和指数性降低流量，所以方法 110 类似于方法 100 且提供许多相同的好处。因此，方法 110 可确保通过子管道的通信量能自动、渐进地达到稳定性。而且，每个管道流量的提高和降低将依赖于管道的最小流量和管道以前的流量。因此，不同管道可以有不同的流量级别，或服务。结果是，方法 110 可产生稳定的行为，公平分配额外带宽并提供分类服务。然而，注意，尽管方法 110 公平分配额外带宽，但控制通过子管道的流量和管道中的拥塞时可以

不考虑公平分配资源。为了为不同子管道或客户提供不同的服务级别，所有网络管理员或其他用户都应该做的是，依赖于某一客户想要的服务水平，为客户设置不同的最小流量，如果需要的话，还有最大流量。因此，方法 110 可通过想要为不同客户或不同媒体提供不同服务级别的因特网服务提供商在多种应用中使用，如使用 DiffServ 的网络。

另外，方法 100 和 110 可以仅在使用方法 100 或 110 控制的网络边缘使用。例如，对网络 10'，方法 100 和 110 可以仅在交换机 14'，也就是管道的入口中执行。因此，计算上相对昂贵的子管道流量控制，不必要在整个网络 10' 中执行。然而，网络 10' 中的管道和子管道通常通过整个网络 10' 定义。因此，通过控制仅在边缘的子管道流量，可确保整个网络 10' 的相应管道不拥塞。从而，不用执行冗余计算和整个网络的子管道控制就能控制拥塞。然而，其它控制方法（最好更简单）也可在网络的相同交换机和其它交换机中结合方法 100 和 110 一起使用。例如，在共同待审并且转让给本申请的受让人的 PCT 专利申请号 No.GB00/04410，标题为“在计算机网络中控制分组传输的方法和系统”中描述了一种控制流量的方法。在上面提及的共同待审申请中描述的在管道中控制流量的方法可与本发明一起使用。因此，至少可提供两个控制级别，子管道的较好级别，和子管道的一般级别。

可使用多种机制确定在图 6、7A 和 7B 中分别描述的方法 100 和 110 的步骤 104 和 116 中是否存在拥塞。图 8-11 描述了用这种方法确定某一管道中是否存在拥塞的实施例。然而，在图 8-11 中描述的方法可用于多个管道，可被组合使用，且可使用其它方法（未示出）。另外，图 8-11 描述的方法可用于拥塞的第二个定义，该定义可用于方法 110 的步骤 120 中，以确定额外带宽是否存在。

图 8 描述了为确定拥塞是否存在的方法 140 的一个实施例。方法 140 利用 ECN。ECN 在 IETF 建议 RFC2481 中描述，该建议适用如 TCP 的协议。在 ECN 中，分组中两个未使用位被用来指示拥塞。支持 ECN 且分组从中经过的交换机设置首位。分组从中经过的支持 ECN 的交换机设置第二位，前提是支持 ECN 的交换机拥塞。因此，位组合

指示分组流经其中的子管道的至少一个交换机是否拥塞。这些位典型在分组从源到目的地流动时设置。当接收方（目的地）发回一个确认时，保存这两个位的合成，并提供给源（发送者）。因此，由这两位代表的 ECN 可被用来确定管道是否拥塞。

特别地，通过步骤 142，确定 ECN 指示经过拥塞交换机的分组数的度量。在一个实施例中，步骤 142 包括确定流经 ECN 指示经过一个拥塞交换机的管道的分组片断。如上讨论，该片断可被用作 PCL。然后通过步骤 144 确定拥塞是否存在（如为 ECN 定义的那样）。在一个实施例中，步骤 144 包括确定 ECN 指示经过一个拥塞交换机的分组片断是否大于某个阈值。然而，一些其它的使用 ECN 的拥塞统计度量可被使用。如果确定没有超过阈值，那么，通过步骤 146，管道将被定义为不拥塞。如果确定超过阈值，那么，通过步骤 148，管道将被定义为拥塞。因此，使用方法 140 可确定管道是否拥塞。

图 9 描述了用于确定管道是否拥塞的方法 150 的另一个实施例。方法 150 利用同步（SYN）分组。在 TCP 中，SYN 分组可指示两个部件，如主机之间会话的开始。当会话分组没有到达目的地时，例如由于拥塞分组被丢弃，部件之间将开始一个新的会话。从而，发出一个新的 SYN 分组。由于 SYN 分组可用于度量会话开始和重新开始的次数，所以它们可被用于在管道中度量拥塞。

通过步骤 152，确定管道中 SYN 分组数目的度量。在一个实施例中，步骤 152 包括确定 SYN 分组与流经管道分组的全部分组数目的片断。该片断可被设置为上面讨论的 PCL。通过步骤 154，然后确定为 SYN 分组定义的拥塞是否存在。在一个实施例中，步骤 154 包括确定 SYN 分组与分组全部数目的比值是否大于某个阈值。然而，一些其它使用 SYN 分组的拥塞统计度量可被使用。如果确定没有超过阈值，那么，通过步骤 156，管道将被定义为不拥塞。如果确定超过阈值，那么，通过步骤 158，管道将被定义为拥塞。因此，使用方法 150 可确定管道是否拥塞。

图 10 描述了用于确定管道是否拥塞的方法 160 的另一个实施例。

方法 160 利用分组在源（发送者）和目的地（接收者）间传送并且将已收到分组的通知返回源的往返时间（RTT）。当一个会话的分组没有到达目的地时，例如由于拥塞而被丢弃，或当分组花费较长时间到达目的地时，RTT 将较长。较长的 RTT 指示管道中较多的拥塞。因此，RTT 的长度可用于度量管道中的拥塞。另外，可以使用从发送者到接收者的时间，而不是 RTT。

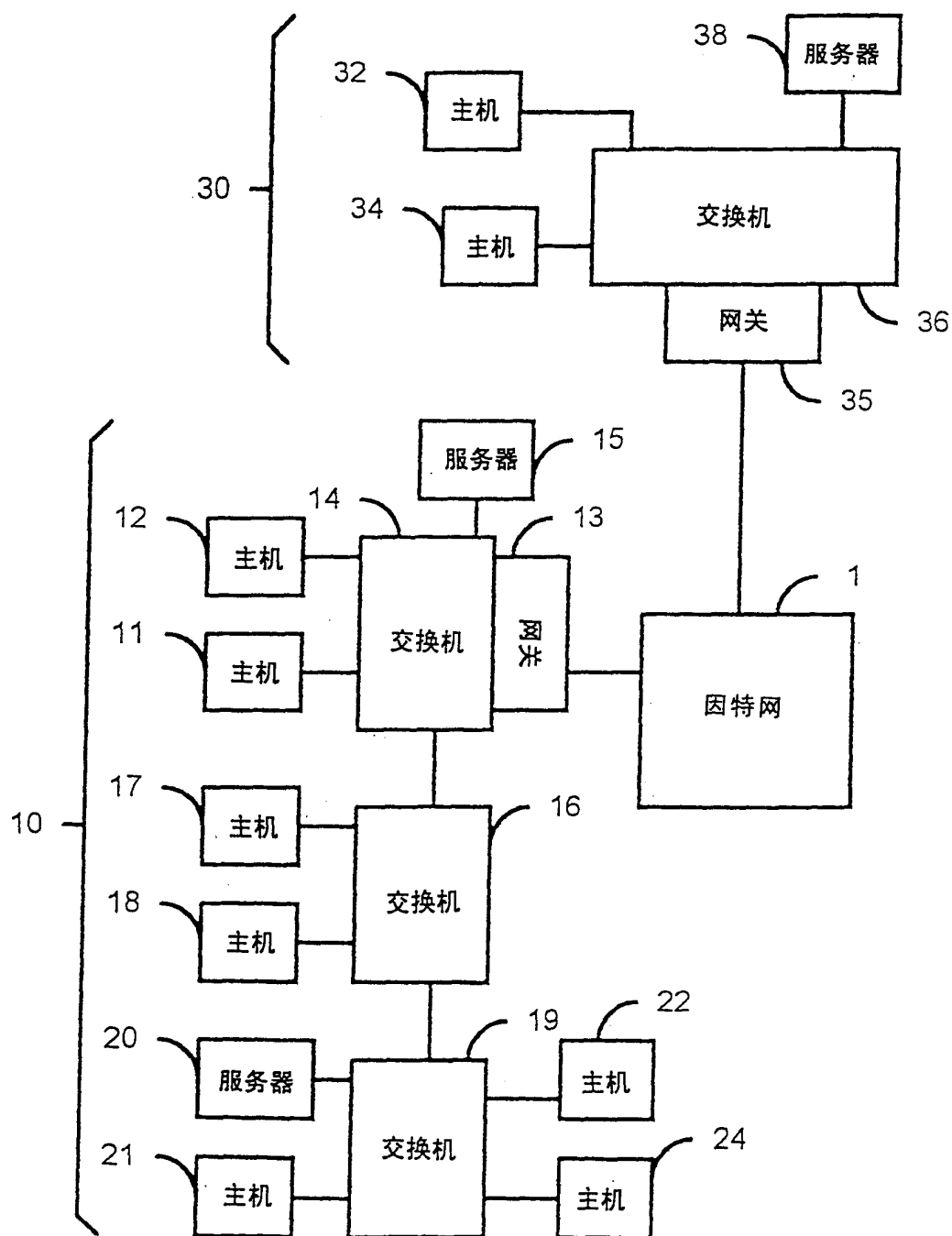
通过步骤 162，可确定管道中分组 RTT 的一些统计度量。在一个实施例中，步骤 162 可包括确定其 RTT 长于某个平均值的分组的片断。然而，也可使用其它 RTT 统计度量。如上讨论，RTT 统计度量可用于确定 PCL。通过步骤 164，例如通过确定 RTT 度量是否在阈值之上，确定为 RTT 定义的拥塞是否存在。如果确定 RTT 度量指示管道不拥塞，则通过步骤 166，管道被定义为不拥塞。如果确定 RTT 度量指示管道拥塞，则通过步骤 168，管道被定义为拥塞。因此，使用方法 160 可确定管道是否拥塞。

图 11 描述了方法 170 的一个实施例，其中 ECN、SYN 分组和 RTT 被结合使用以确定管道是否拥塞。通过步骤 172，确定 ECN 指示经过拥塞交换机的分组数目的度量。在一个实施例中，步骤 172 包括确定 ECN 指示经过拥塞交换机的分组的片断。通过步骤 174，确定管道中 SYN 分组数目的度量。在一个实施例中，步骤 174 包括确定 SYN 分组与流经管道分组的全部数目的片断。通过步骤 176，确定一些管道中分组的 RTT 的统计度量。通过步骤 178，组合 ECN、SYN 分组和 RTT 的度量，以提供拥塞的度量，如 PCL。通过步骤 180，最好通过确定 PCL 是否超过阈值，来确定是否存在拥塞。如果确定 PCL 指示管道不拥塞，那么通过步骤 182，定义管道不拥塞。如果确定 PCL 指示管道拥塞，那么通过步骤 184，定义管道拥塞。因此，使用方法 170 可确定管道是否拥塞。

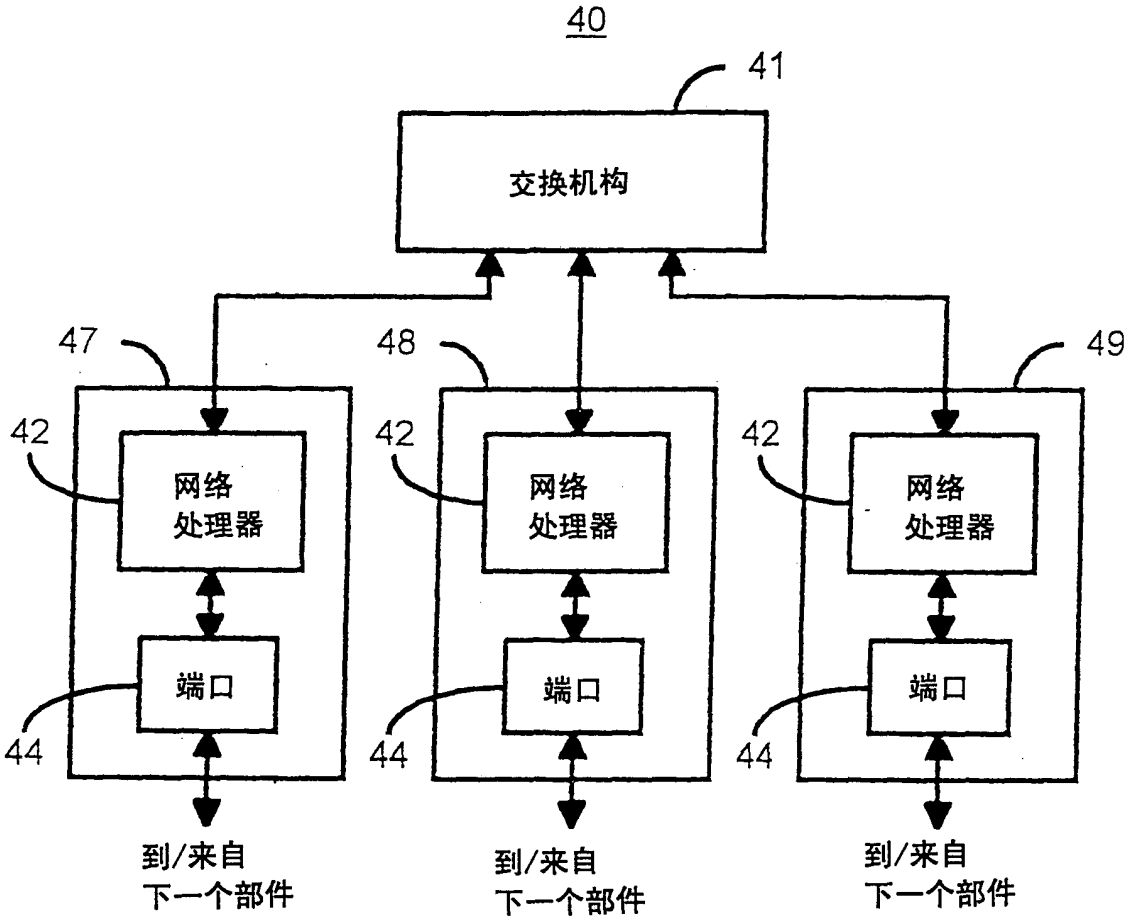
因此，可使用多种方法确定管道是否拥塞。如果管道拥塞，则控制管道内子管道的流量。否则，在子管道级别不控制流量。子管道的流量通过指数性降低流量和，在一些情况下，线性提高流量而得到控

制。因此，仅当由于拥塞存在，网络运行不理想的时候，才消耗资源在子管道级别精细地控制流量。而且，拥塞也可用于决定怎样精细地调整子管道流量，例如通过用于确定额外带宽是否存在的拥塞的第二个定义。另外，可为不同子管道，因此还有不同管道，提供不同的服务级别。而且，仅在网络边缘提供控制级别。从而，使子管道中流量的冗余控制变得不必要。然而，也可以使用网络中控制流量的其它方法。

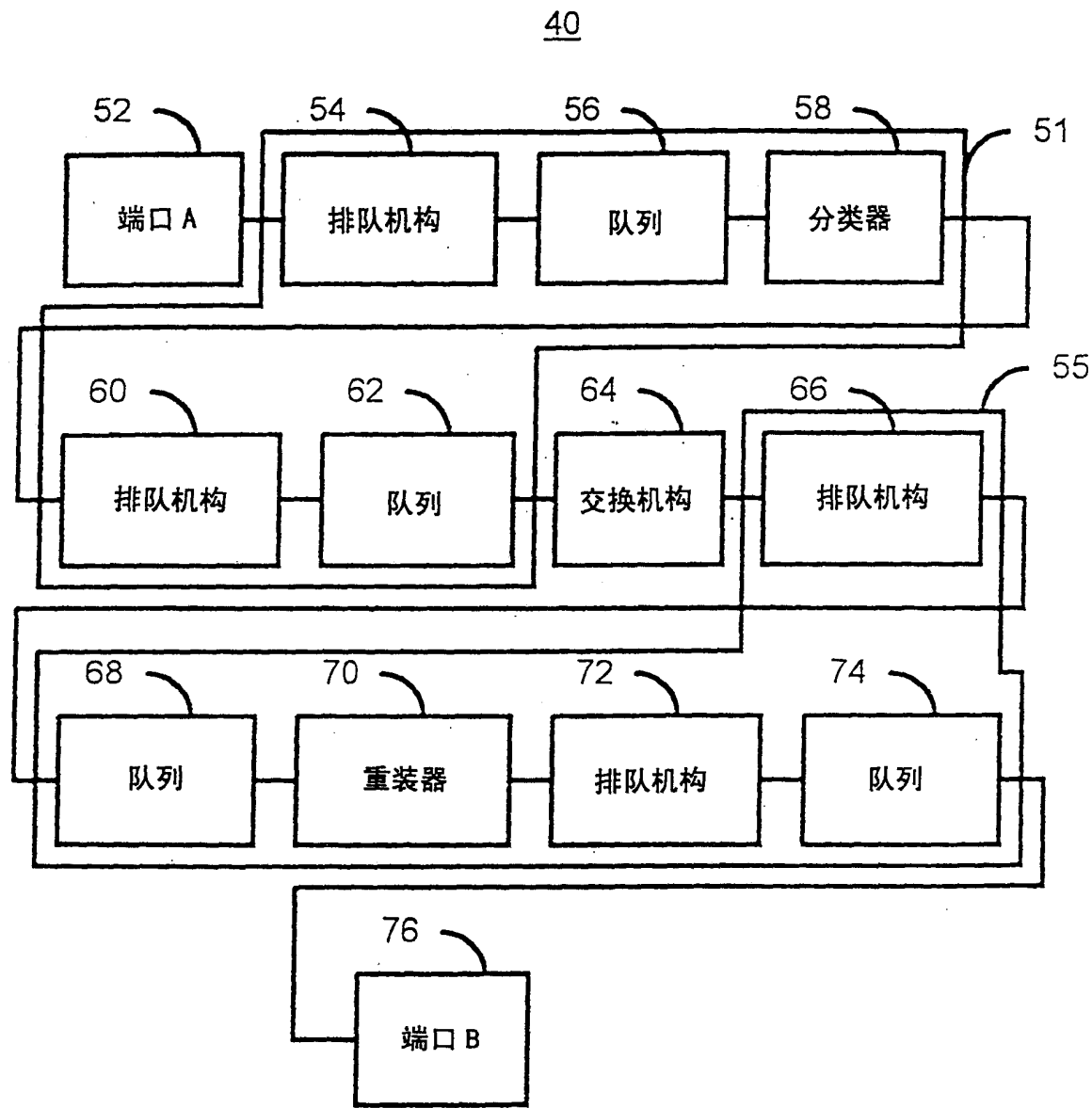
公开了为控制网络的子管道中的流量的方法和系统。根据本发明写的软件保存在计算机可读的一些介质中，如存储器、CD-ROM 或通过网络传输且由处理器执行的介质。



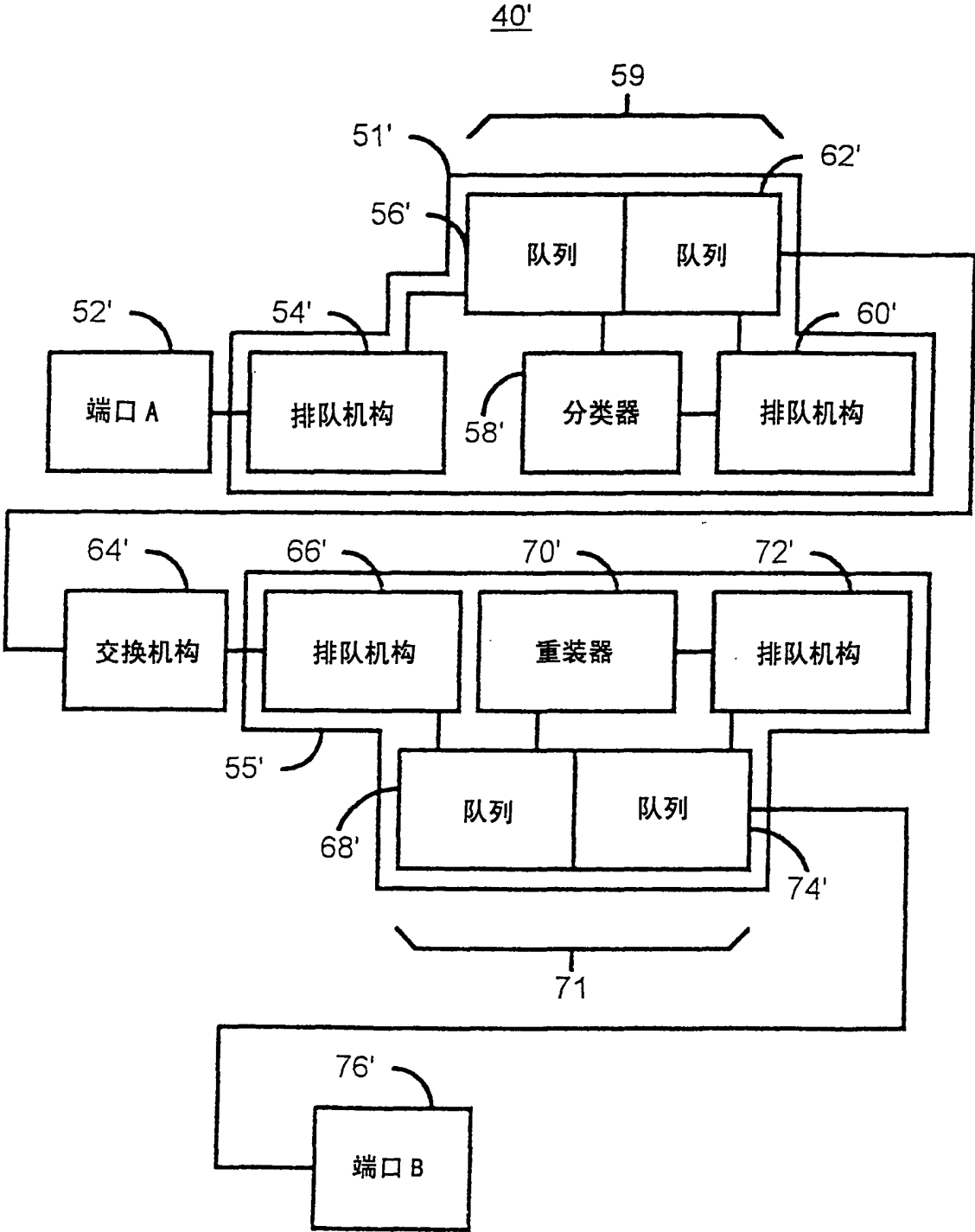
现有技术
图 1



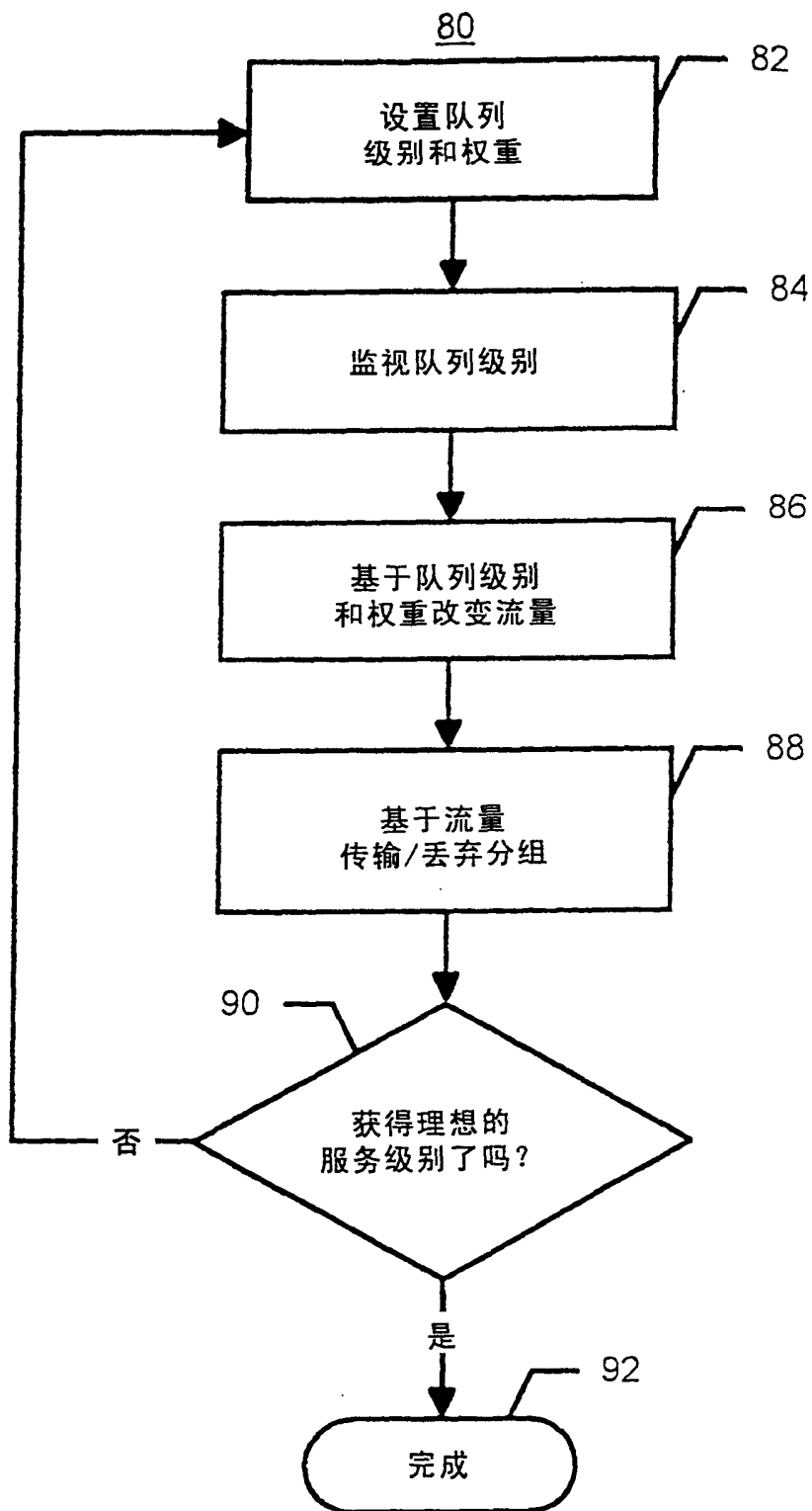
现有技术
图 2



现有技术
图 3A



现有技术
图 3B



现有技术
图 4

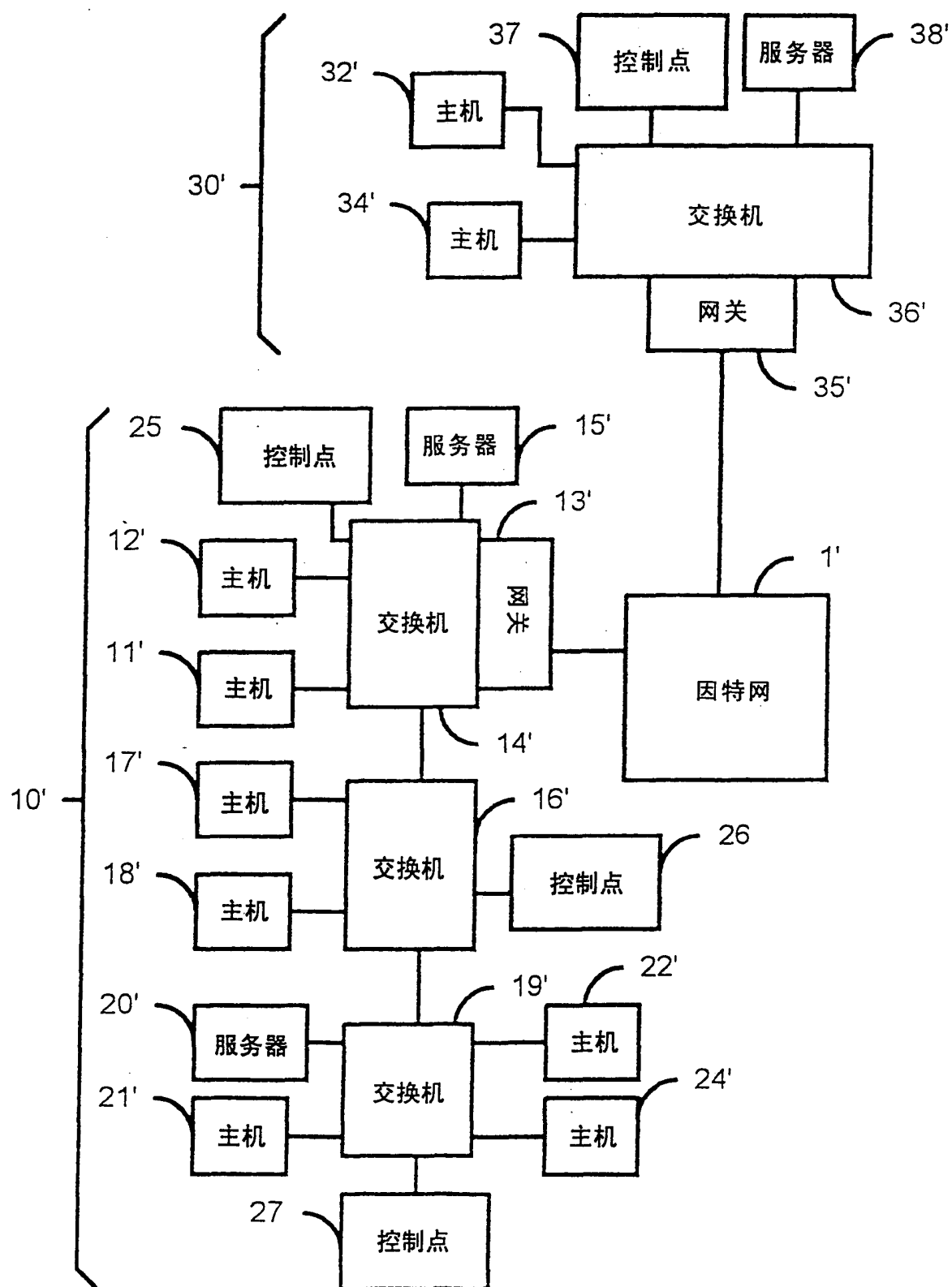


图 5

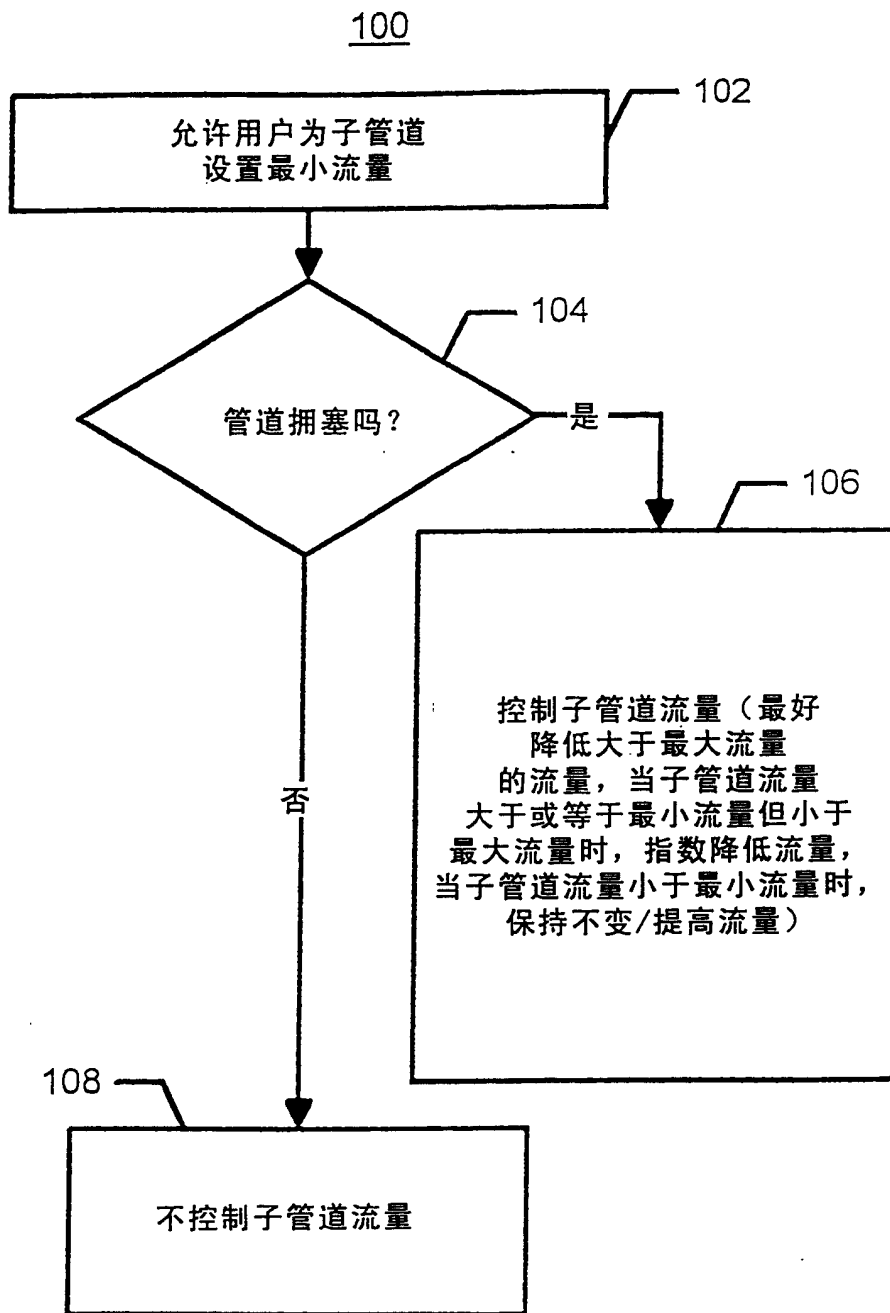


图 6

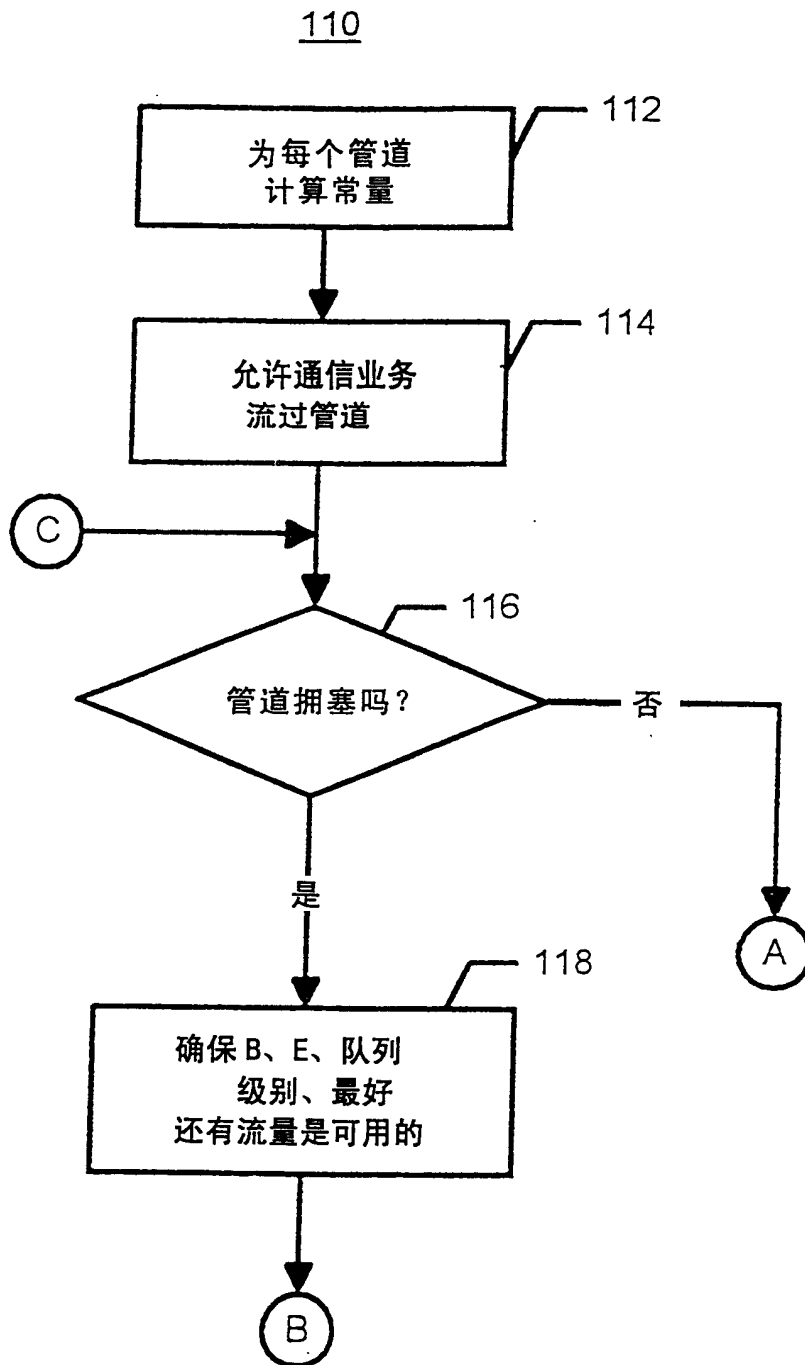


图 7A

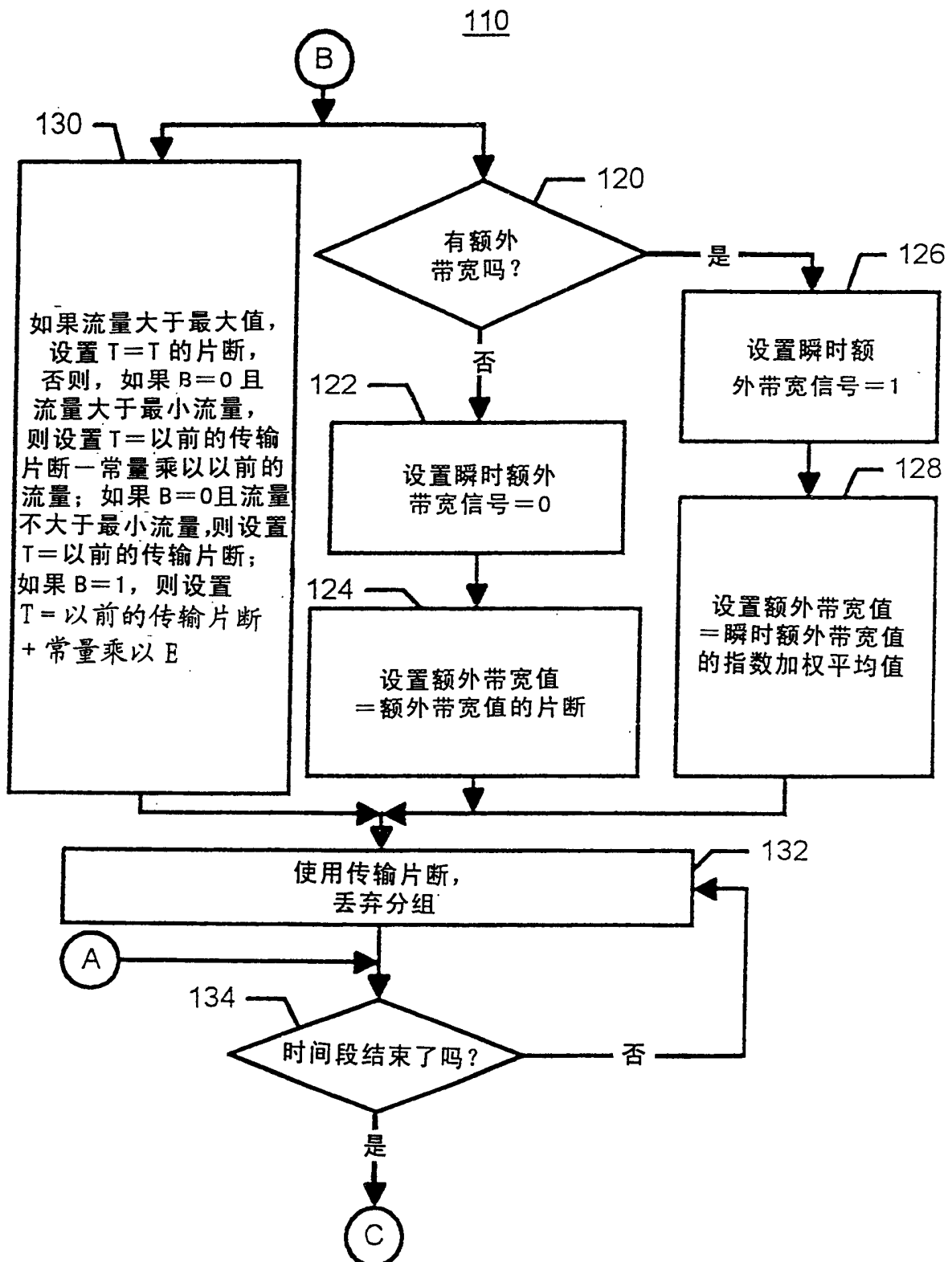


图 7B

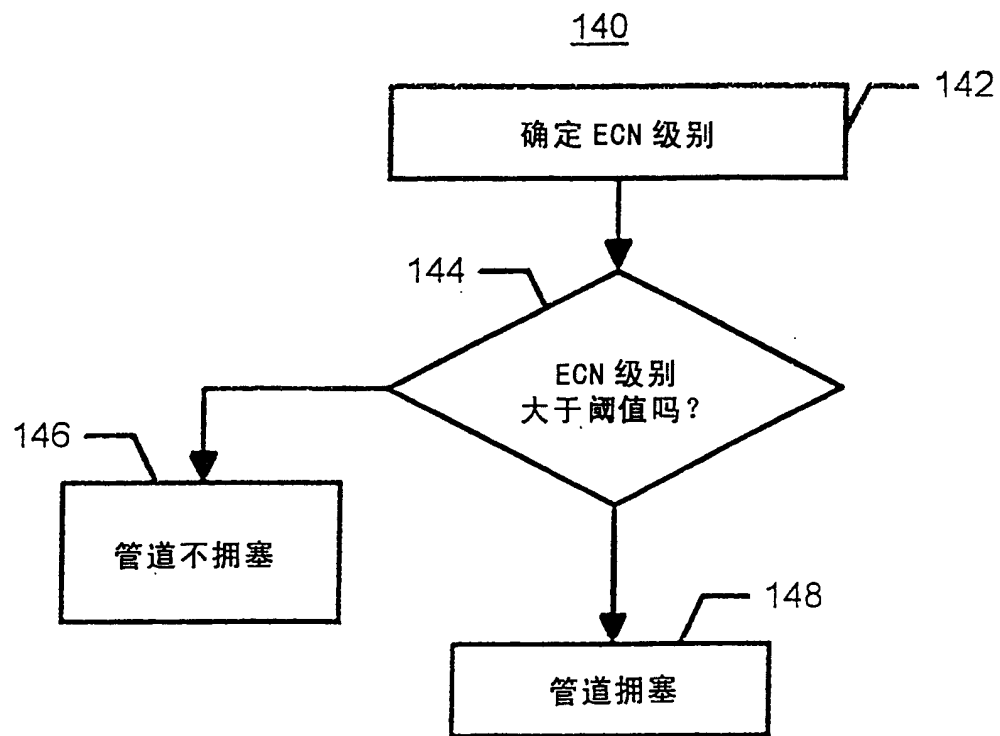


图 8

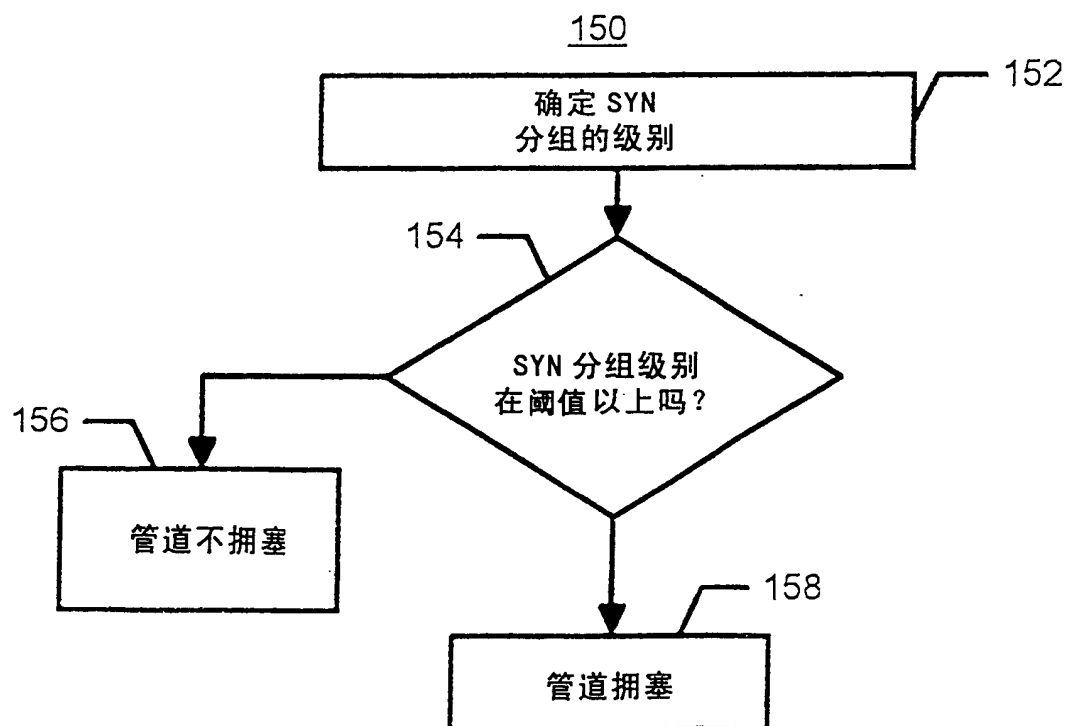


图 9

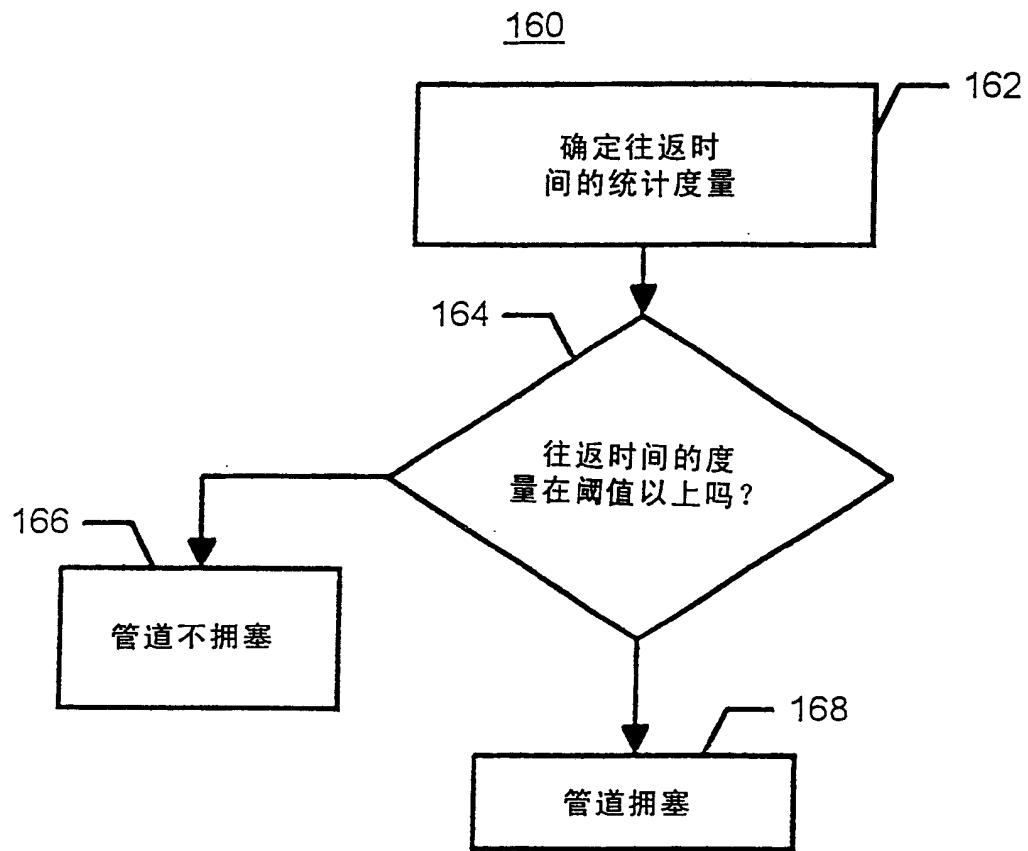


图 10

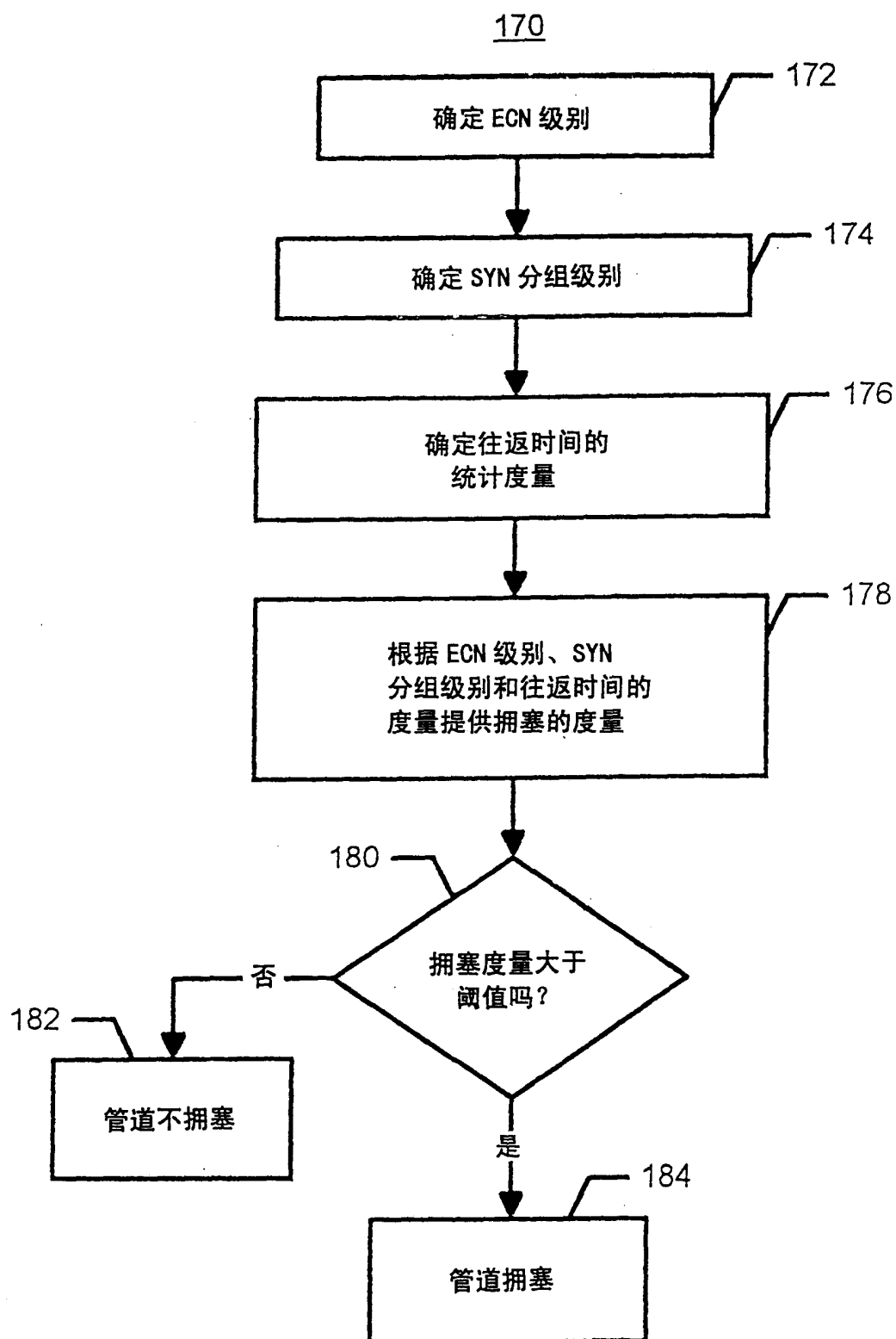


图 11