

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H04L 12/56 (2006.01)



[12] 发明专利说明书

专利号 ZL 03815866.3

[45] 授权公告日 2007 年 5 月 16 日

[11] 授权公告号 CN 1316802C

[22] 申请日 2003.5.8 [21] 申请号 03815866.3

[30] 优先权

[32] 2002. 5. 29 [33] US [31] 10/158,291

[86] 国际申请 PCT/US2003/015729 2003. 5. 8

[87] 国际公布 WO2003/103236 英 2003. 12. 11

[85] 进入国家阶段日期 2005. 1. 4

[73] 专利权人 英特尔公司

地址 美国加利福尼亚州

[72] 发明人 A·黄 J·-M·开亚 J·凌

J·-C·卡尔德龙 V·约斯

[56] 参考文献

CN1151641A 1997. 6. 11

WO01/63858A1 2001. 8. 30

US 5867663 A 1999. 2. 2

EO0706298A2 1996. 4. 10

审查员 李 楠

[74] 专利代理机构 上海专利商标事务所有限公司

代理人 张 鑫

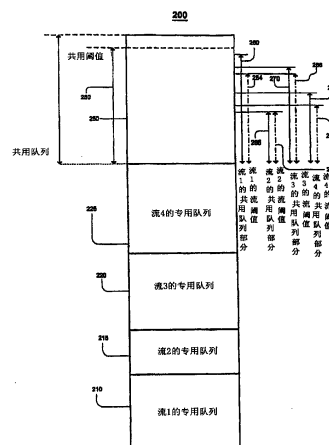
权利要求书 4 页 说明书 5 页 附图 6 页

[54] 发明名称

缓冲存储器保存法

[57] 摘要

网络应用要求担保的吞吐速率，其建立方法是用缓冲存储器保存法管理用来存贮输入分组的数据队列。缓冲存储器保存法把一部分数据队列保存为各流的专用队列，把另一部分数据队列保存为共用队列，并使一部分共用队列与各流相关联。专用队列规模与共用队列部分规模对所有的流所保存的缓冲存储量的总量，可以超过缓冲输入分组可用的物理存储器的总量。



1. 一种分组交换应用的缓冲存储器管理方法，其特征在于，所述方法包括：
使多个分组流的每一个都与专用队列、和共用队列的特定部分相关联，
其中每个专用队列具有专用队列规模，并且共用队列的每个特定部分具有规模和流阈值，
其中共用队列具有共用阈值，并且
其中所有的流的专用队列规模和共用队列规模的总和超过缓冲分组可用的物理存储器总量；
若与特定流关联的专用队列的未使用部分规模大于或等于特定分组规模，把来自特定分组流的特定分组接入与该特定流关联的专用队列；
若特定分组不被与特定流关联的专用队列接受、与特定流关联的共用队列部分的未使用部分规模大于或等于该特定分组的规模、而且共用队列的使用部分规模小于或等于共用阈值，就把来自特定分组流的特定分组接入相关联的共用队列的特定部分；
以及
如果：
a) 特定分组不被与特定流关联的专用队列接受，且
b) 与特定流关联的共用队列使用部分规模大于共用队列的特定部分的流阈值，
或者所有使用的共用队列的特定部分的规模超过共用阈值，
则丢掉来自特定分组流的特定分组。
2. 如权利要求 1 所述的方法，其特征在于，所述专用队列的规模对不同的流不同。
3. 如权利要求 1 所述的方法，其特征在于，所述专用队列的规模对所有的流都相同。
4. 如权利要求 1 所述的方法，其特征在于，所述方法还包括：
使各收到的分组与被接入共用队列的概率相关联，
若该特定分组不被与特定流关联的专用队列接受，而且与特定分组关联的概率大于与一个或多个未被与接收分组关联的专用队列接受的其它接收分组关联的概率，就把来自特定分组流的特定分组接入共用队列，和
若特定分组未接入与特定流关联的专用队列，而且该特定分组未被接入共用队列，就丢掉来自特定分组流的特定分组。

5. 如权利要求 4 所述的方法，其特征在于，共用阈值小于共用队列规模，所述方法还包括：

如果：

特定分组不被与特定流关联的专用队列接受，
共用队列使用部分的规模大于共用阈值，而且
共用队列的特定部分的使用部分规模小于或等于特定部分的流阈值，

就使来自特定分组流的特定分组与第一概率相关联；和

如果：

特定分组不被与特定流关联的专用队列接受，
共用队列使用部分的规模小于或等于共用阈值，而且，
共用队列的特定部分的使用部分规模大于特定部分的流阈值，

则使来自特定分组流的特定分组与第二概率相关联；

其中第一概率小于第二概率。

6. 一种分组交换应用中的缓冲存储器管理设备，该设备包括处理器和连接该处理器的存储器，其特征在于，所述处理器包括一个或多个元件，用于：

使多个分组流的每个分组流都与专用队列、和共用队列的特定部分相关联，

其中每个专用队列具有专用队列规模，并且共用队列的每个特定部分具有规模和流阈值，

其中共用队列具有共用阈值，并且

其中所有的流的专用队列规模和共用队列规模的总和超过缓冲分组可用的物理存储器总量，和

若专用队列与特定流关联的未使用部分的规模大于或等于该特定分组的规模，就把来自该特定分组流的特定分组接入与特定流关联的专用队列，

若特定分组不被与特定流关联的专用队列接受、与特定流关联的共用队列部分的未使用部分规模大于或等于该特定分组的规模、而且共用队列的使用部分规模小于或等于共用阈值，就把来自特定分组流的特定分组接入相关联的共用队列的特定部分；以及

如果：

a) 特定分组不被与特定流关联的专用队列接受，且

b) 与特定流关联的共用队列使用部分规模大于共用队列的特定部分的流阈值，

或者所有使用的共用队列的特定部分的规模超过共用阈值，

则丢掉来自特定分组流的特定分组。

7. 如权利要求 6 所述的设备，其特征在于，所述专用队列的规模对不同的流不同。

8. 如权利要求 6 所述的设备, 其特征在于, 所述专用队列的规模对所有的流都相同。

9. 如权利要求 6 所述的设备, 其特征在于, 所述处理器还包括一个或多个元件, 用于:

使各收到的分组与被接入共用队列的概率相关联,

若该特定分组不被与特定流关联的专用队列接受, 而且与特定分组关联的概率大于与一个或多个未被与接收分组关联的专用队列接受的其它接收分组关联的概率, 就把来自特定分组流的特定分组接入共用队列, 和

若特定分组未接入与特定流关联的专用队列, 而且该特定分组未被接入共用队列, 就丢掉来自特定分组流的特定分组。

10. 如权利要求 9 所述的设备, 其特征在于, 所述共用阈值小于共用队列规模, 处理器还包括一个或多个元件, 用于:

如果:

特定分组不被与特定流关联的专用队列接受,

共用队列的使用部分的规模大于共用阈值, 而且

共用队列的特定部分的使用部分规模小于或等于特定部分的流阈值,

就使来自特定分组流的特定分组与第一概率相关联; 和

如果:

特定分组不被与特定流关联的专用队列接受,

共用队列的使用部分的规模小于或等于共用阈值, 而且,

共用队列的特定部分的使用部分规模大于特定部分的流阈值,

则使来自特定分组流的特定分组与第二概率相关联;

其中第一概率小于第二概率。

11. 一种分组交换应用中的缓冲存储器管理系统, 其特征在于, 所述系统包括:

业务管理设备;

耦接传输信道的端口; 和

业务管理设备与端口间的链路,

其中, 业务管理设备包括一个或多个元件, 用于:

使多个分组流的每个分组流都与专用队列和共用队列特定部分相关联,

其中每个专用队列具有专用队列规模, 并且共用队列的每个特定部分具有规模和流阈值,

其中共用队列具有共用阈值, 并且

其中所有的流的专用队列规模和共用队列规模的总和超过缓冲该分组可用的物理存储器总量，和

若与特定流关联的专用队列的未使用部分的规模大于或等于特定分组规模，把来自特定分组流的特定分组接入与该特定流关联的专用队列，

若特定分组不被与特定流关联的专用队列接受、与特定流关联的共用队列部分的未使用部分规模大于或等于该特定分组的规模、而且共用队列的使用部分规模小于或等于共用阈值，就把来自特定分组流的特定分组接入相关联的共用队列的特定部分；以及

如果：

a) 特定分组不被与特定流关联的专用队列接受，且

b) 与特定流关联的共用队列使用部分规模大于共用队列的特定部分的流阈值，或者所有使用的共用队列的特定部分的规模超过共用阈值，

则丢掉来自特定分组流的特定分组。

缓冲存储器保存法

背景

以下描述涉及数字通信系统，尤其涉及包括传输分组的高速分组交换网的系统。诸如异步转移模式（ATM）、网间协议（IP）与千兆以太网等高速分组交换网，都支持大量对输入分组争夺缓冲存储器空间的不同预约时间（session）的连接。

数字通信系统一般应用发送称为分组的数据块的分组交换系统。通常，待发送的消息或其它数据组都大于分组的规模，必须被分成一连串分组，各分组包含一部分被发送数据和首部内用于将该分组通过网络送到其目的地的控制信息。

附图简介

图 1 是分组交换网视图。

图 2 是用来存贮输入分组的缓冲存储器视图。

图 3 是保存缓冲存储器空间以存贮输入分组的执行过程的状态图。

图 4~6 是说明保留缓冲存储器空间以存贮输入分组的过程的流程图。

图中以同样的标号表示同样的单元。

详细描述

图 1 示出一典型的分组交换系统，包括经通信路径 115 连接分组交换网 120 的发送服务器 110，而分组交换网 120 经通信路径 125 接目的地服务器 130。发送服务器 110 通过分组交换网 120 向目的地服务器 130 发送作为一连串分组的消息。在分组交换网 120 中，分组一般通过一连串服务器。各分组到达服务器时，服务器先将该分组暂时存缓冲存储器，然后把分组发送给下一服务器。分组继续通过该网，直至到达目的地服务器 130，而后者收到该分组时，就把它暂时存入缓冲存储器 135。

高速分组交换网能支持大量连接（也称为流），如有些宽带网可在每个线路卡里通过 64 个逻辑端口支持 256,000 条连接。来自某一流的各输入分组在接收后被存入缓冲存储器里的数据队列，若没有存贮特定分组的缓冲存储器空间，该输入分组就被丢掉。

网络应用要求经担保的吞吐速率，这可用管理存贮输入分组的数据队列的缓冲存

储器的保存法建立。缓冲存储器保存法把一部分数据队列保存为各流的专用队列，把另一部分数据队列保存为共用队列，并使一部分共用队列与各流相关联。对各流设置的专用队列规模为输入分组提供了担保的吞吐速率，而在最大速率超过担保吞吐速率的周期内，共用队列提供了缓冲分组空间。专用队列有一个或多个分配给它的保留存储器地址，或者与特定的保留存储器地址无关被分配存储器空间。同样地，共用队列有一个或多个分配给它的保留存储器地址，或与特定的保留存储器地址无关被分配存储器空间。专用队列与共用队列部分对所有的流所保存的缓冲存储器总量，可以超过缓冲分组可用的物理存储器的总量，但是专用队列部分所保留的缓冲存储器总量可以不超过缓冲分组可用的物理存储器的总量。

如图 2 所示，用来存贮输入分组的数据队列 200 的缓冲存储器被分成各流专用的队列 210、215、220 与 225 和共用队列 250。为简便起见，图 2 只示出一小部分数据队列 200。与各流关联的共用队列 250 部分用箭头 260、265、270、275 示出，80%的共用队列规模与部分 260 里的第一流关联，40%的共用队列规模与部分 265 里的第二流关联，75%的共用队列规模与部分 270 里的第三流关联，而 55%的共用队列规模与部分 275 里的第四流关联。专用队列 210、215、220、225 和共用队列部分 260、265、270、275 的规模之和，超过存储输入分组可用的物理存储器的现有总量。

在从确定数据队列中有存贮特定输入分组的空间时到存贮该特定输入分组时的时段内，会减少数据队列 200 的未使用部分，这种数据队列未使用部分的减少会妨碍存贮该特定输入分组，造成输入分组丢掉。

小于共用队列规模的共用阈值 280，可减少因数据队列增大而被丢掉的输入分组数量。共用阈值 280 可以设为小于或等于共用队列 250 规模的值，实际阈值的选用基于丢失分组似然性（随共用阈值一起增大）与共用队列使用效率（随共用阈值一起减少）之间的平衡。此外，还对各流设置一流阈值 284~287，它小于或等于与该流关联的共用队列部分 260、265、270、275 的规模。

缓冲存储器保存实施法所用的专用队列规模，可以对所有的流一样或者对各流不一样。某一种实施法可使用对所有的流一样的流阈值，也可改变各流的流阈值，或者不用流阈值。

图 3 示出在处理器上执行缓冲存储器保留的状态图 300。在接收了输入分组后，处理器可以在与一流关联的专用队列里存贮来自该流的输入分组（状态 310），可以在共用队列里存贮该输入分组（状态 320），或者可以丢掉该分组（状态 330）。

若专用队列里对分组有空间（过渡 342、344、346），处理器就在与流关联的专用队列里存贮来自该流的输入分组（状态 310）。对一特定的流而言，处理器保持状态 310（过渡 342），直到该流的专用队列占满。

在专用队列里无空间时（过渡 348），若该流的共用队列部分和共用队列里有空间（过渡 350），就把输入分组存贮在共用队列里（状态 320）。流的共用队列部分和共用队列都必须有空间，因为共用队列可用的物理存储器小于分配给所有流的共用

队列部分之和的空间总量。若共用队列或专用队列里没有存贮输入分组的空间（过渡 354、356），输入分组就从分组流中丢掉（状态 330）。处理器继续丢掉输入分组，直到共用队列（过渡 352）或专用队列（过渡 346）里变成有空间。

参照图 4，过程 400 用输入分组的规模判断共用队列部分里对流是否有空间。用一共用队列的共用阈值实施图 4 的过程 400，该阈值等于共用队列的规模，而且不使流阈值与接收输入分组的流关联。

过程 400 在处理器从流中接收输入分组时开始（410），处理器判断流的专用队列规模的未使用部分是否大于或等于分组规模（420），若是，处理器就在该流的专用队列里存贮该分组（430），并等待接收流的另一个输入分组（410）。

若处理器断定专用队列规模的未使用部分小于分组规模（如在流专用队列中没有存贮该分组的空间），则处理器判断流的共用队列部分的未使用部分规模是否大于或等于该分组（440），若不是，就丢掉该分组（450）。因为流的专用队列与共用队列部分都没有足够的存贮分组空间，只好丢掉该分组。分组丢掉后，处理器等待接收另一输入分组（410）。

若处理器断定该流的共用队列部分的未使用部分规模大于或等于分组规模，处理器就判断共用队列的使用部分是否小于或等于共用阈值（460），若是，处理器在共用队列中存贮该分组（470），并等待接收流的另一输入分组（410）。若处理器断定其用队列规模的使用部分大于共用阈值，处理器则丢掉该分组（450），等待接收流的输入分组（410）。

参照图 5，过程 500 用一流阈值判断流的共用队列部分里是否有空间。过程 500 对该共用队列使用小于共用队列规模的共用阈值，并与流阈值小于同该流关联的共用队列部分规模的各流相关联。

过程 500 在处理器从流接收输入分组时开始（510），判断流的专用队列是否有适合该分组的空间（520），当有空间时，在该流的专用队列里存贮进入分组（530）；若在该流的专用队列中没有空间（520），处理器判断共用队列部分的使用部分是否小于或等于流阈值（540）。这与对图 4 描述的实施情况相反，其中处理器根据输入分组的规模来判断共用队列部分是否有空间，并不使用流阈值。

若符合该流阈值，处理器就判断共用队列的使用部分是否小于或等于共用阈值（550）。只有在共用队列部分的使用部分和共用队列小于或等于它们各自的阈值时，处理器才在共用队列中存贮该分组（560），否则处理器就丢掉该输入分组（570）。然后，处理器等待输入分组（510），继续上述过程。

参照图 6，过程 600 对收到的特定分组分配一允许接入共用队列的概率，当该特定分组的被接受概率高于分配给其它争用缓冲存储器空间的输入分组的概率时，就允许该收到的分组进入共用队列。

过程 600 在处理器从流中接收输入分组时开始（610），判断该流的专用队列是否有适合该分组的空间（620），而在有空间时，在该流的专用队列里存贮该输入分

组（630）；若在该流的专用队列里没有存贮分组空间，则处理器判断该流共用队列部分的使用部分是否小于或等于该流阈值（640），判断该共用队列的使用部分是否小于或等于该共用阈值（650）。根据这些判断，处理器可能丢掉该分组，或如下表所列那样在共用队列中存贮该分组。

共用队列部分的使用部分小于或等于流阈值	共用队列的使用部分小于或等于共用阈值	对分组分配概率（选项）	存贮结果
是	是		在共用队列里存贮分组
是	否	对分组分配较高概率	若分组概率高于争用的分组，在共用队列存贮分组，否则丢掉分组
否	是	对分组分配较低概率	若分组概率高于争用分组，在共用队列里存贮分组，否则丢掉分组
否	否		丢掉分组

当共用队列部分的使用部分大于流阈值，而且共用队列的使用部分大于共用阈值时，分组被丢掉（660）。

当共用队列部分的使用部分小于或等于流阈值，而且共用队列规模的使用部分小于或等于共用阈值时，在共用队列里存贮分组（670）。

若这两个条件都不存在，则在共用队列部分的使用部分小于或等于流阈值，而且共用队列的使用部分大于共用阈值时，处理器对该分组分配在共用队列里被存贮的较高概率（680）。当共用队列部分的使用部分大于流阈值时，而且共用队列的使用部分小于或等于共用阈值时，处理器对该分组分配在共用队列里被存贮的较低概率（685）。然后，处理器判断分配给该分组的概率是否大于分配给争用缓冲存储器空间的其它输入分组的概率（690）；若是，处理器在共用队列里存贮分组（670），否则丢掉分组（660）。

缓冲存储器保存法为输入分组提供担保的吞吐速率，避免缓冲器拥塞。缓冲存储器保存技术提供各种管理网络应用的参数，包括共用阈值、各流的流阈值、各流的专用队列、共用队列和各流的共用队列部分。有些实施法预先指定该参数，而另一些实施法可根据当前网络条件来改变诸参数。

适合分组应用的缓冲存储器保存法的诸优点适用于应用定长或可变长分组的其

它分组交换网实施法。

实施法包括方法或过程、设备或系统或者计算机媒体上的计算机软件。应该理解，可作出各种修改而不背离下述如权利要求的精神与范围，例如若以不同的次序执行所揭示技术的诸步骤和/或揭示的系统中的诸元件以不同方式组合和/或被其它元件置换或增补，仍可得到有利的结果。

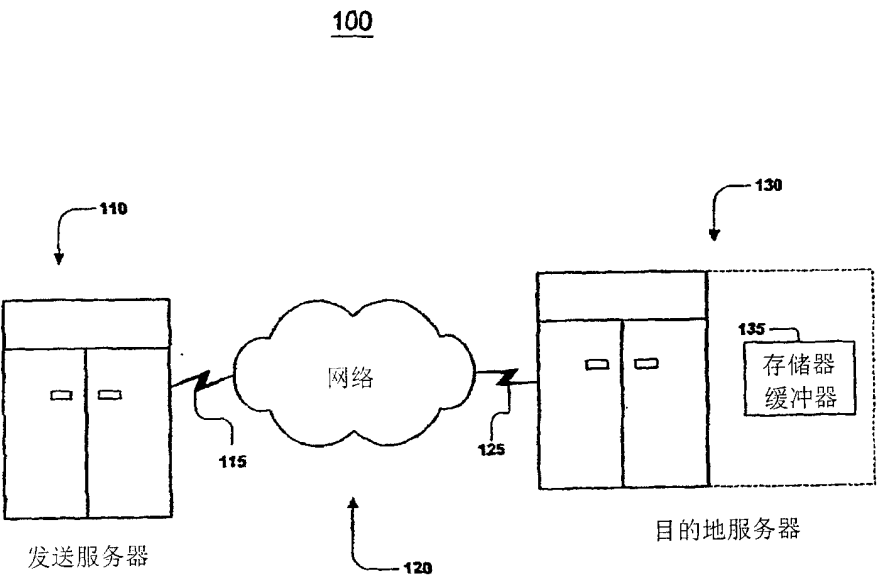


图 1

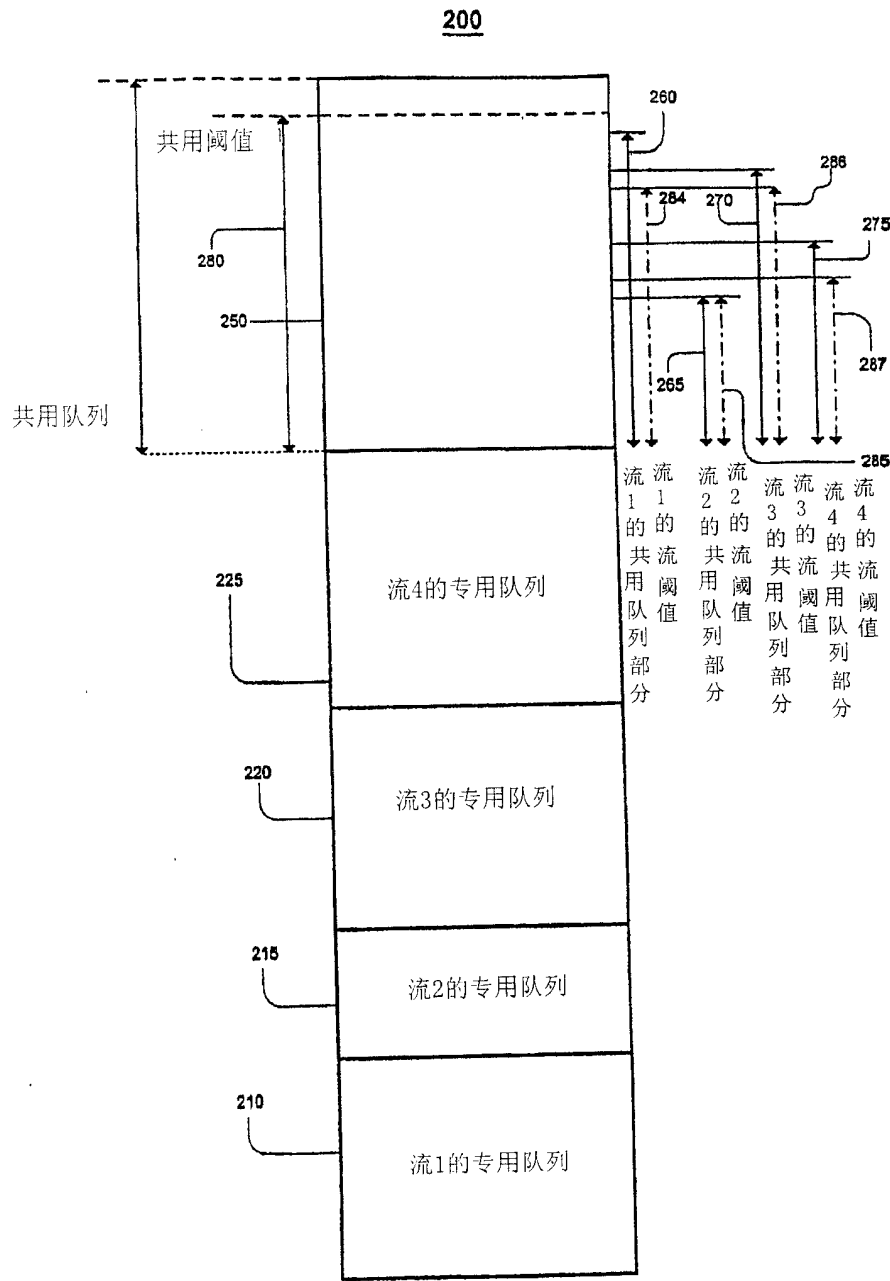


图 2

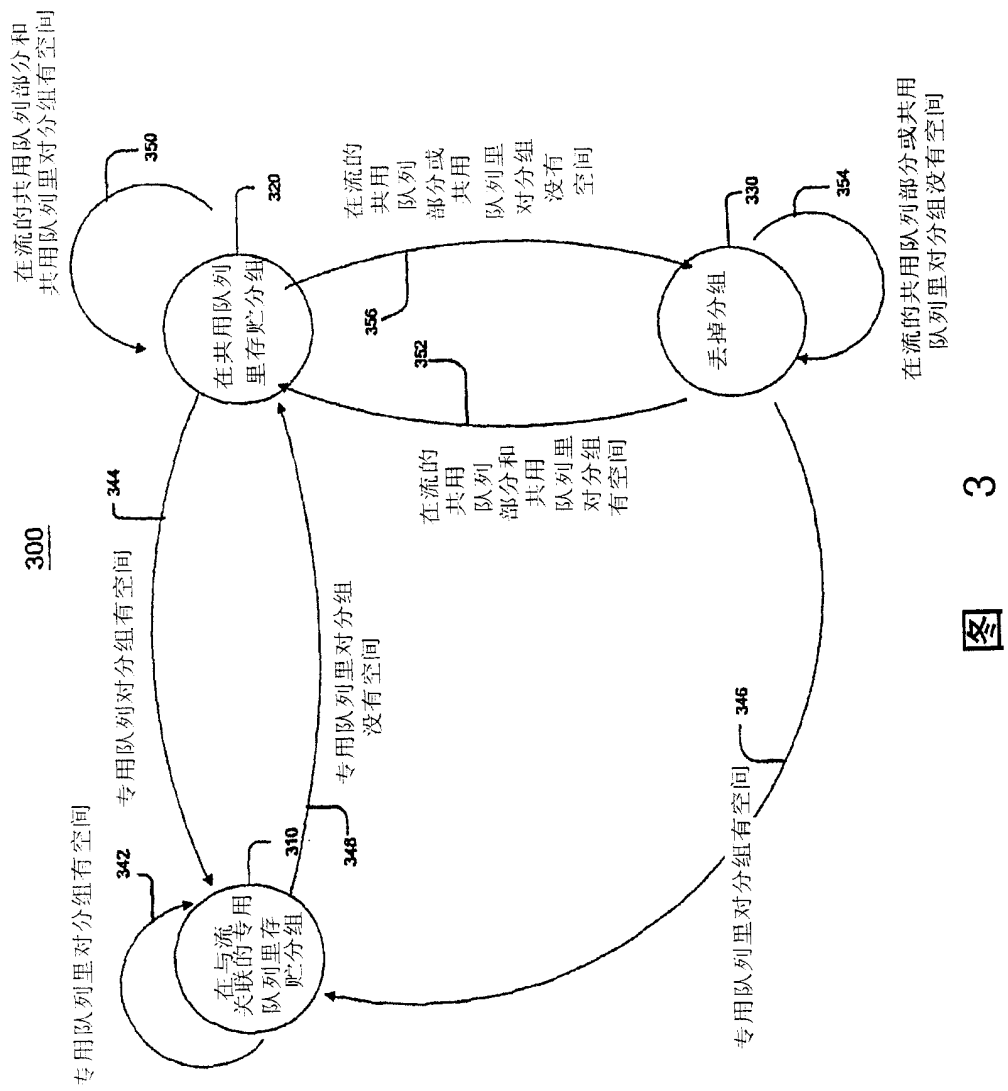


图 3

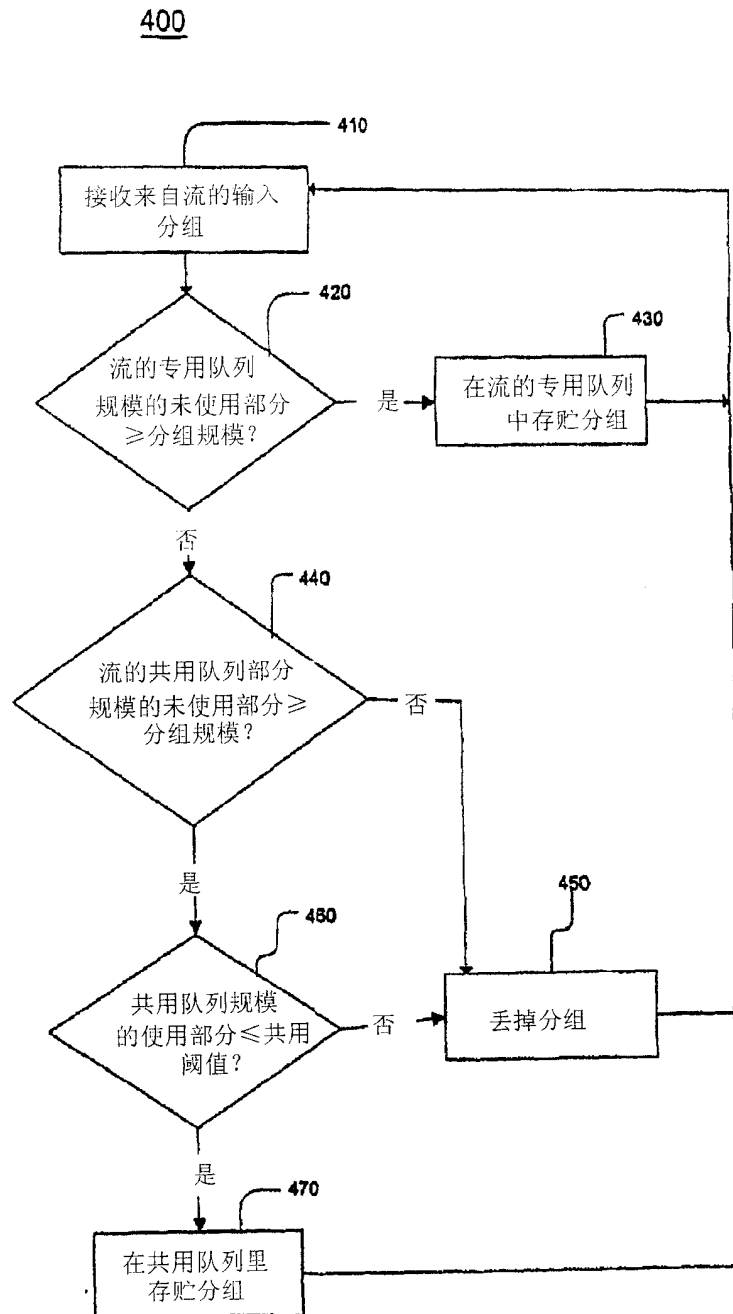


图 4

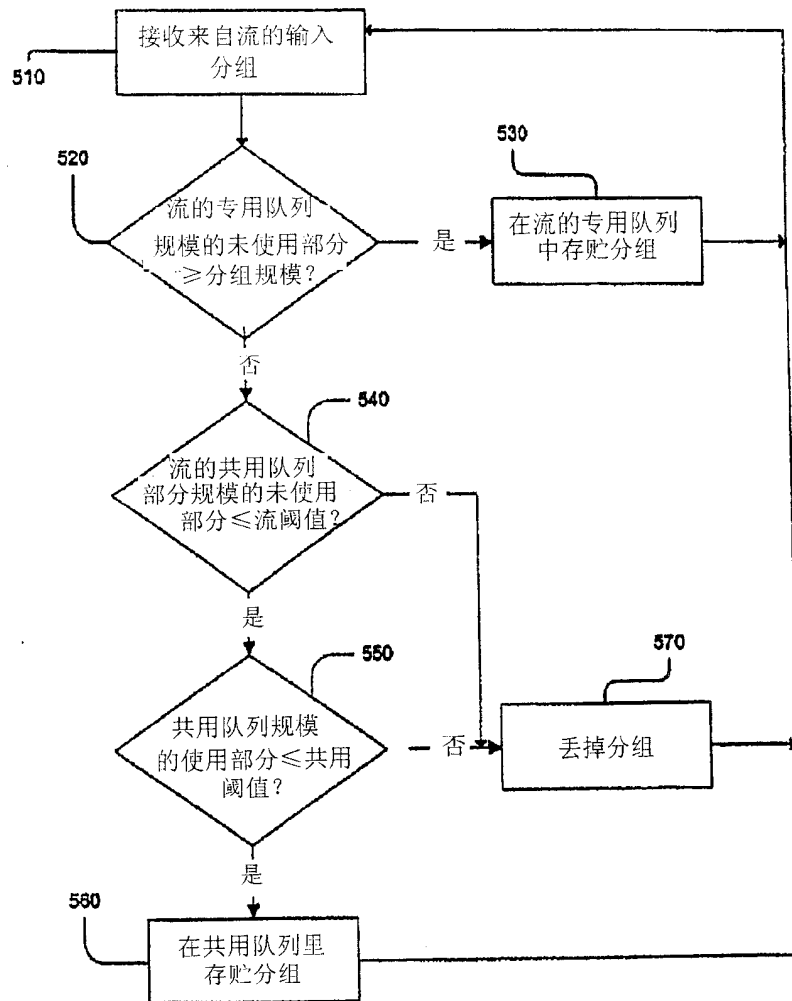
500

图 5

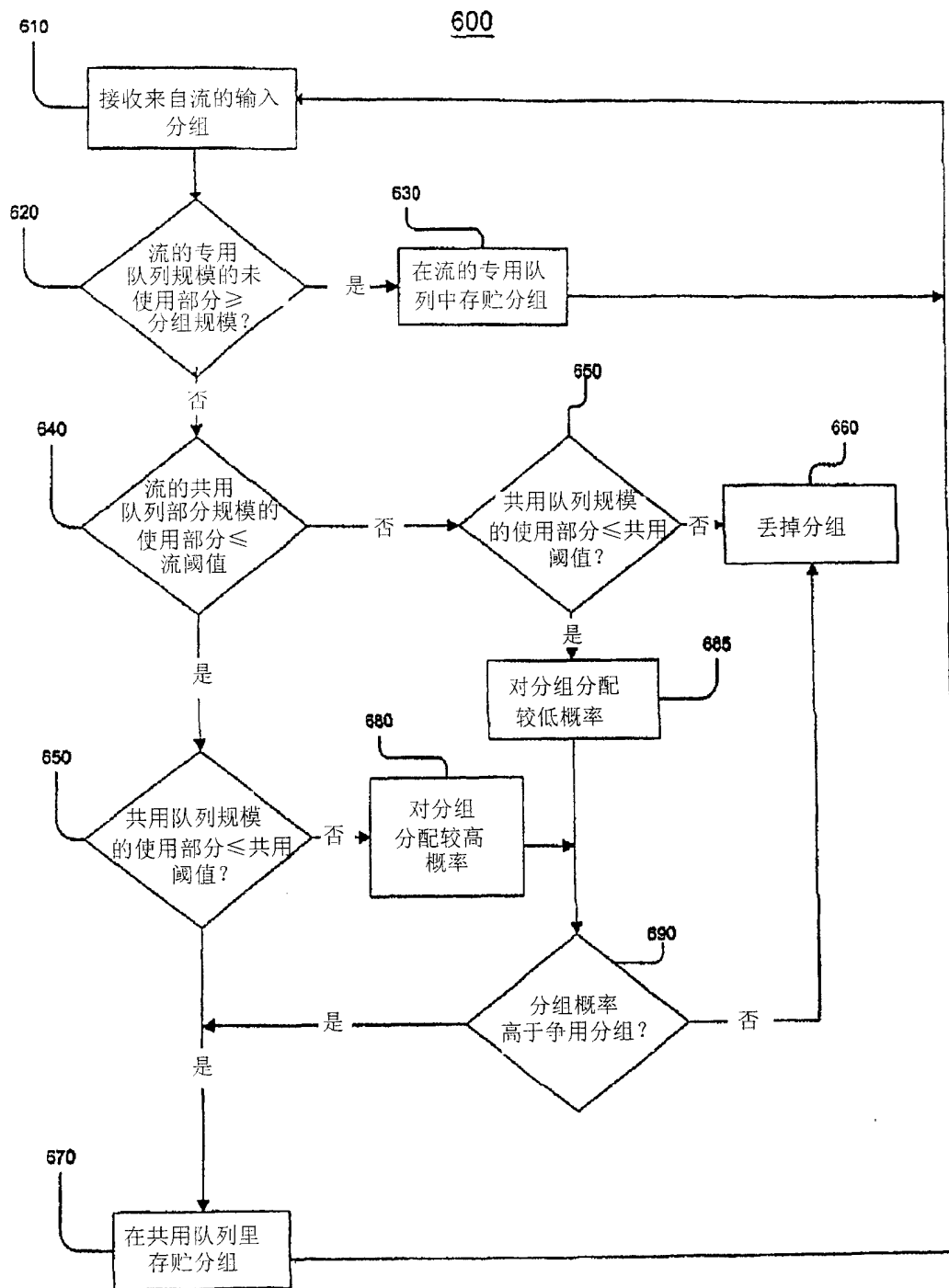


图 6