



[12] 发明专利申请公开说明书

[21] 申请号 200310102317.8

[43] 公开日 2004年5月26日

[11] 公开号 CN 1499786A

[22] 申请日 2003.10.24

[21] 申请号 200310102317.8

[30] 优先权

[32] 2002.10.25 [33] US [31] 10/280, 959

[71] 申请人 阿尔卡特公司

地址 法国巴黎

[72] 发明人 默德哈法·哈桑-阿里

杰夫·门德尔松 安妮·拉斯泰罗

陈立胜 雷迪米尔·希尔施图特

赛纳·索尔托尼

弗朗西斯科·莫雷诺

[74] 专利代理机构 北京市金杜律师事务所

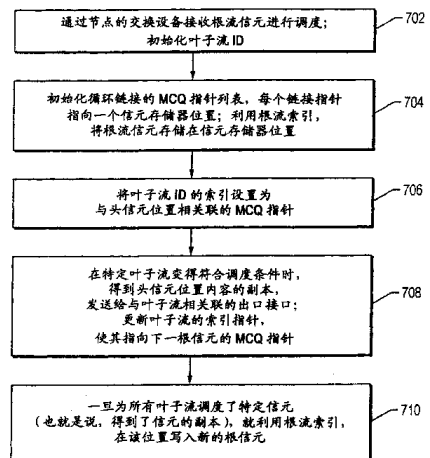
代理人 张维

权利要求书 2 页 说明书 17 页 附图 7 页

[54] 发明名称 接入节点的 ATM 交换设备中使用的组播系统和方法

[57] 摘要

本发明公开了一种调度 ATM 环境中的根流的组播方案。包括索引存储器结构的缓冲器系统从入口接口得到多个与叶子流关联的叶子流索引值，以及与根流相关联的根流索引。提供了一种指针存储器，其中多个循环链接的指针根据叶子和根流进行索引。提供了一种信元存储器结构，它拥有链接的指针指向的多个信元存储器位置。ATM 环境中的根流信元根据根流索引存储在信元存储器位置中。当特定流变得符合调度条件时，得到一个根信元副本进行发送，该根信元位于某个链接指针所指向的信元存储器位置，链接指针指向该特定叶子流。



I S S N 1 0 0 8 - 4 2 7 4

1. 一种引导 ATM 环境中的根流的组播方法，使用对应数量的叶子流将 ATM 环境中的根流引导向多个出口接口，包括以下步骤：

5 初始化循环链接的指针表，其中多个链接的指针指向存储空间中对应数量的信元存储器位置；

 在上述 ATM 环境中接收到根流信元时，基于与所述根流相关联的根流索引，将所述信元存储在所述信元存储器位置，其中所述根流索引初始化成所述链接指针链表的特定指针，用于存储头根流信元，在附加根流信元到达时，通过所述链接指针更新所述根流索引；

10 初始化多个叶子流索引值，其中每个索引值对应于一个叶子流，索引值初始化成指向包含所述头根流信元的信元存储器位置的所述特定指针；

 当特定叶子流变得符合调度条件时，得到一个所述头根信元副本，发送给与所述特定叶子流相关联的出口接口，更新其叶子流索引值，指向所述链接指针表中的下一项；以及

 在为每个所述叶子流调度特定根流信元时，将新的根流信元写入所述特定根流信元的存储器位置。

2. 根据权利要求 1 所述的引导 ATM 环境中的根流的组播方法，其中确定特定叶子流变得符合调度条件的所述操作，由具有多个聚合层的分层调度器完成，这些聚合层在多个业务优先级类间划分。

3. 根据权利要求 2 所述的引导 ATM 环境中的根流的组播方法，其中所述多个业务优先级类与恒定比特率（CBR）业务类别相关联。

4. 根据权利要求 2 所述的引导 ATM 环境中的根流的组播方法，其中所述多个业务优先级类与实时可变比特率（rt-VBR）业务类别相关联。

5. 根据权利要求 2 所述的引导 ATM 环境中的根流的组播方法，其中所述多个业务优先级类与非实时可变比特率（nrt-VBR）业务类别相关联。

-
6. 根据权利要求2所述的引导 ATM 环境中的根流的组播方法，其中所述多个业务优先级类与未指定比特率（UBR）业务类别相关联。
7. 根据权利要求2所述的引导 ATM 环境中的根流的组播方法，其中所述多个业务优先级类与保证帧频（GFR）业务类别相关联。
8. 根据权利要求1所述的引导 ATM 环境中的根流的组播方法，其中每个叶子流由一个流 ID（FID）指定。
9. 根据权利要求1所述的引导 ATM 环境中的根流的组播方法，还包括增加附加数量的叶子流来接收所述根流的步骤。
10. 根据权利要求1所述的引导 ATM 环境中的根流的组播方法，还包括删除一定数量的已有叶子流的步骤。

接入节点的 ATM 交换设备中使用的组播系统和方法

5 相关申请

本申请公开的主题涉及下列共有的共同未决的美国专利申请所公开的主题： (i) James W. Dove 等人的“Stackplane Architecture”，于 1999 年 12 月 22 日申请，申请号为 09/469,897; (ii) Eric Friedrichs 等人的“Scalable Architecture For An Access Node”，于 2002 年 6 月 27 日申请，申请号为 10/280,604； (代理卷号为: 1285-0090US); (iii) Thornton Collins 等人的“Integrated Gateway Functionality In An Access Network Element”，于 2001 年 11 月 2 日申请，申请号为 10/052,846; (iv) Mudhafar Hassan-Ali 等人的“System And Method For Implementing GFR Service In An Access Node’s ATM Switch Fabric,”，于同一天申请的，美国申请号为 10/280,700; (代理卷号为: 1285-0102US); (v) Mudhafar Hassan-Ali 等人的“Virtual Group Connection Scheme For ATM Architecture In An Access Node”，于同一天申请的，美国申请号为 10/280,604; (代理卷号为: 1285-0099US); (vi) Mudhafar Hassan-Ali 等人的“Calendar Heap System And Method For Efficient Sorting,”，于同一天申请的，美国申请号为 10/281,033 ； (代理卷号为: 1285-0101US); (vii) Mudhafar Hassan-Ali 等人的“Hierarchical Scheduler Architecture For Use With An Access Node,”，于同一天申请的，美国申请号为 10/280,894 ； (代理卷号为: 1285-0103US); 这些专利申请通过引用并入本发明。

25 技术领域

本发明总的来说涉及电信领域。确切地说，本发明涉及在接入节点的异步传输模式 (ATM) 交换设备中实现组播方法的一种系统和方法，但本发明并不局限于此。

背景技术

远程接入市场正在经历一个大的转变。三个因素促成了这种转变。第一是用户数量的增长，例如需要高性能因特网的和多媒体远程接入的小型办公/居家办公（SOHO）用户数量的增长。在电信方面自由化的政府行为是另一因素，各地通过消除本地市场规范来培育更为广泛的竞争。第三因素，也是最后一个因素，是公共电话交换网（PSTN）的拥塞，PSTN的设计和开发原本只是为了语音业务量。

电信技术的若干重大进展使得电信网络的骨干连接能够具备高吞吐量。例如，通过在同步光网（SONET）/同步数字体系（SDH）物理层上实现异步传输模式（ATM）网络技术，电信网络能够达到几百兆比特每秒（Mbps）的数据速率。但是，为满足远程接入的带宽需求所做出的努力受限于电信局中心局（CO）和用户的远端站点之间已有的双绞铜线基础设施（也就是接入网），一般称作本地环路。在电信领域中，这些限制有时统一称作“最后一公里”问题。

为避免最后一公里问题所产生的瓶颈，当前接入网解决方案在本地环路中也采用光纤技术。与利用高速电信网一样，基于光纤的本地环路基础设施的体系结构一般采用 SONET 作为物理层技术。除了网络设计的改进，随着光部件和相关光电子的最近进展，使得宽带接入日趋普及。

此外，伴随着因特网用户数量的显著增长，产生了对用分组交换网（PSN）基础设施（例如，那些基于因特网协议（IP）进行寻址的基础设施）来替代当前电信网所采用的已有电路交换网络（CSN）基础设施的浓厚兴趣。从网络运营商角度来讲，分组交换基础设施中固有的业务总量使得每个最终用户的传输成本和基础设施成本得以降低。最终，这种成本的降低使得网络运营商能够将由此节省的成本转移到最终用户身上。

因此，正在开发一种新型的以业务为中心的网络（不同于已有的以话音为中心和数据为中心的网络），用以实现众所周知的下一代网

络 (NGN) 基础设施, 在下一代网络基础设施中综合的语音/数据/视频应用可以利用端到端传输路径中 PSN 上的分组传输机制来提供。前面间接提到, 人们相信, 在接入网中采用分组网络基础设施能够提供更高的传输效率, 更低的操作和运营成本, 以及统一的接入。

传统接入系统允许接入数字本地话音交换机, 例如第 5 类交换机, 这通过延伸多根金属环路, 并将它们集中成一束, 从而高效传送时分复用 (TDM) 话音业务量来实现。一般情况下, 这种接入网络的体系结构使用多种配置下的一个或多个接入节点, 前述配置可以是例如点对点链、环等, 其中接入节点自身可以包括多个信道组, 后者提供的线路接口服务于大量用户。

但是, 为了提供更好的功能和业务提供, 要求当前的接入网支持先进的传输技术, 例如 SONET, 对于节点的内部结构也是如此。在这些节点中, ATM 用于承载除传统的 TDM 业务, 如 T1 和 TDM-DS3 业务之外的大部分用户业务量。因此, 接入节点设计需要支持 TDM 和 ATM 交换设备。

ATM 论坛提供了一组规范, 用以管理 ATM 交换设备不同方面, 包括连接的类型, 例如虚信道连接和虚通道连接, 以及它们的拓扑, 例如点到点 (单点播送) 连接和点到多点 (组播) 连接。众所周知, 在环境中支持组播时, 单个业务量源 (也就是根) 发送信元或分组给多个接收复制业务量的目的地 (也就是叶)。对容错力强的实现而言, 叶子流必须不干扰或妨碍到其它叶子的传输。此外, 需要支持组播流的队列在存储器使用方面应当高效。

虽然已有若干技术能够在 ATM 环境中实现组播, 但它们都具有特定缺陷。首先, 当前组播方案需要大量使用存储器, 因为每个叶子的队列都通过不同的信元缓冲器来实现。此外, 叶子的性能是相互关联的, 因为单个叶子节点因某中原因无法工作时, 该节点会导致组播处理中止。

发明内容

因此，本发明提供了一种在 ATM 环境中调度根流的组播方案，前述 ATM 环境可以是例如交换设备，其中多个出口接口由相应数量的叶子流提供服务，这些叶子流具有最小缓冲器需求，叶子流之间的分离性得到改进。包括索引存储器结构的缓冲器系统从入口接口得到多个与叶子流关联的叶子流索引值，以及与根流相关联的根流索引。提供了一种指针存储器，其中多个循环链接的指针根据叶子和根流进行索引。提供了一种信元存储器结构，它拥有链接的指针指向的多个信元存储器位置。ATM 环境中的根流信元根据根流索引存储在信元存储器位置中。当特定流变得符合调度条件时，得到一个根信元副本进行发送，该根信元位于某个链接指针所指向的信元存储器位置，链接指针指向该特定叶子流。

一方面，本发明涉及一种在异步传输模式（ATM）环境中的组播方法，用于将 ATM 根流通过对应数量的叶子流引导向多个出口接口，包括以下步骤：初始化循环链接的指针表，其中多个链接的指针指向存储空间中对应数量的信元存储器位置；在所述 ATM 环境中接收到根流信元时，基于与所述根流相关联的根流索引，将所述信元存储在所述信元存储器位置，其中所述根流索引初始化成所述链接指针链表的特定指针，用于存储头根流信元，在附加根流信元到达时，通过所述链接指针更新所述根流索引；初始化多个叶子流索引值，其中每个索引值对应于一个叶子流，索引值初始化成指向包含所述头根流信元的信元存储器位置的所述特定指针；当特定叶子流变得符合调度条件时，得到一个所述头根信元副本，发送给与所述特定叶子流相关联的出口接口，更新其叶子流索引值，指向所述链接指针表中的下一项；以及，在为每个所述叶子流调度特定根流信元时，将新的根流信元写入所述特定根流信元的存储器位置。

另一方面，本发明涉及一种组播系统，用于将 ATM 环境中，例如接入节点中的交换设备的根流通过对应数量的叶子流导向多个出口接口，其中提供了各种缓冲器结构及相关的调度器方案，用以完

成前述操作。

附图说明

通过下面的详细描述,并结合附图,可以对本发明有更为全面的
5 理解,在附图中:

图 1 给出了示例性的接入节点,该节点的 ATM 交换设备中可以
有利地实现本发明的教导;

图 2 给出了多层多平面调度器的一种实施方式,它在图 1 所示
ATM 交换设备中实现 GFR 业务;

10 图 3 描述的高层功能框图说明了实现本发明教导的调度器漏桶
模块和优先级队列模块;

图 4 描述了一种示例性组播流,其中按照本发明的教导将根流发
送给多个出口接口(也就是叶子);

图 5 给出了实现本发明组播队列缓冲器系统的存储器实施例;

15 图 6 给出了一种循环链接指针系统的实施例,用于为具有 8 个信
元存储器位置深度的组播队列缓冲器提供服务;以及

图 7 给出了本发明组播方法的一种实施例中采用的操作流程图。

具体实施方式

20 下面根据 Mudhafar Hassan-Ali 等人同一天提交的题为“Hierarchical
Scheduler Architecture For Use With An Access Node”的美国专利申
请,申请号为 10/280,894(代理人卷号:1285-0103US)的共同拥
有共同未决美国专利申请(此后称为分层调度器体系结构申请)的
教导,给出本发明的一种实施方式,前述专利申请在此引用并入本
25 发明。在该申请中详细描述到,接入网中的电信节点可以包括一种
可调整的体系结构,其中提供 TDM 和 ATM 交换设备,用以支持性
能水平的提高。

此外,与 ATM 交换设备关联的调度功能可以按照业务类别(也
就是业务平面),并且跨多种层次数据管道集(也就是子端口、总线

级、架级、栈面级以及管道级等等，视为各种聚合层)进行划分，这对可调整硬件体系结构而言是必要的，这样，接入网络节点的 ATM 交换设备中能够有效实现业务量合同遵从，以及必要连接隔离和公平带宽分配。

5 下面参看本专利申请的附图，其中相同或类似的元件在几张图中标以相同的标号，给出的各种元件不一定按比例绘出，尤其是参看图 1，该图描述了一种例示性的接入节点 100，它具有 ATM 交换设备 102 的高层功能性表示，设备 102 可以有利地实现本发明的教导。在前面引用的分层调度器体系结构申请中解释了，交换设备 102 的
10 总体功能包括：监管；运行、管理和维护 (OAM)；信头转换；排队和许可控制；以及调度和业务量整形。可以很容易地看出，发往交换设备 102 的业务量通过多个接口提供。传输接口 104 用于将节点的设备连接到骨干网，例如 ATM 网络 105。栈面接口 106 用于将业务量从辅架组链 107(例如包括分层调度器体系结构申请的图 5 所示
15 信道组 506-1 到 506-4，以及信道组 508-1 到 508-4)传送到交换设备 102。通过线路单元 (LU) 107-1 到 107-N 的多个用户接口举例说明了各种业务源，如 xDSL，T1，ISDN，DS-3/OC-3 等，这些业务源可以通过适当的总线级端口 109-1 到 109-N，与交换设备 102 相连。线路单元接口中的一个接口可以连接到作为接入网 (未在该图中示
20 出)一部分的 RT 111。

针对内部 ATM 业务量可以定义两种类型 ATM 连接：虚信道连接 (VCC) 和虚通道连接 (VPC)。VCC 一般是 ATM 连接能有的最小单元，可以由包括一对物理接口上的标识符，也就是虚信道标识 (VCI) 和虚通道标识 (VPI) 的唯一值来表示。VPC 则定义为共享
25 同一 VPI 值和共同资源池 (例如带宽等)的一组全部的流。因此，可以看出，VP 是一束 VC，这样，通过减少需要管理的部件数量，简化了 ATM 环境下的连接管理，其中每个连接由其唯一 VPI/VCI 对来标识。

从拓扑角度来看，VCC 或 VPC 可以是以下两种类型之一：(i)

点到点连接，其中建立双向连接，且每个方向中的信源可以不同，以及 (ii) 点到多点连接，它通常采用多个单向连接，在设备间实现组播传输。下面通过详细描述将会看到，本发明提供了一种调度 MC 流的有利方案，它利用最少量的存储器，同时将叶子流的性能彼此

5 分离。

此外，也可以在本发明中实现另一层 ATM 连接层次，称为虚拟组连接或 VGC。有关 VGC 实现的细节在 Mudhafar Hassan-Ali 等人同一天提交的题为“Virtual Group Connection Scheme For ATM Architecture In An Access Node”，美国专利申请号：10/280,604；(代理卷号为：1285-0099US)的共同拥有共同未决美国专利中给出，该专利通过引用并入。

10

因此，进入交换设备 102 (其功能可以以 ATM 交叉连接交换设备 (XCF) 卡形式实现) 的流的入口业务量管理一般包括三个阶段：监管、VC 队列/缓冲器分配和整形/调度。在一种实施方式中，与这些级相关的硬件可以整合到 XCF 卡中。监管器的主要功能是确保接收的信元与所应用的连接描述符一致。如果不一致，就可以通过清除/设置信元头中的信元丢弃优先级 (CLP) 来丢弃入信元或者打标记 (也就是带标记)。一般来说，监管功能采用 ITU-T 1.371 和 ATM 论坛 ATMF-TM-121 标准中描述的众所周知的算法来实现。实际上，

15 这些算法 (一般称为通用信元速率算法或 GCRA) 采用所谓的信贷计数器，称为桶 (bucket)，以及信贷，称为令牌。在接收到信元时，如果计数器 (也就是桶) 有足够的信贷 (也就是令牌)，那么接纳该信元；否则，该信元标记成低优先级信元或者被丢弃。此外，在分层调度器体系结构申请中解释了，基于 GCRA 的算法可以在所谓的漏桶模块 (LBM) 中实现，用以实现若干 ATM 业务类别 (基于业务类别 (CoS) 排序，业务类别由类似于时间敏感性、峰值和持续带宽保证、突发性和交付保证等因素定义) 的业务量监管和业务量整形 (也就是调度)。

20

层次调度功能通过优先级队列模块 (PQM) (该图中未示出) 实

现，其中各层对条目实现调度功能，根据流集合，该条目可以是以下之一：子端口、总线、端口和管道。实际上，当交换设备接收到新流的信元时，该数据流由调度器中的一个条目如下表示。从LBM接收到流ID（也就是FID），基于CoS/QoS，流数据（也就是FID和时间戳或TS），存储在应用的第一层数据结构中。在第一层的所有5 的竞争子端口中（例如线路单元的不同流），层仲裁器只会选择一个具有最小TS的子端口，然后将它转发给下一层的仲裁机构，也就是第二层仲裁。第二层数据结构相应包含不同子端口的“优胜FID/TS”数据。同样，只有一个具有最小TS的条目被选出，转发给第三层。10 对其它的聚合层如此重复这种处理，最终得到每种业务优先级类的优胜提名（也就是优胜信元的FID/TS数据）。在分层调度器体系结构申请中还指出了，对每个业务类平面执行这种层基仲裁，从而得到每个平面的优胜提名，这样，考虑CoS的基于时间戳的仲裁器对各个业务平面的优胜者进行仲裁，选择最终的优胜者。

15 因此，应当认识到，各层的仲裁涉及管理与其相关的PQ结构，用于选择该层的优胜者。一般来说，PQ结构以树的形式实现，其中数据节点（代表例如接纳信元或低层选择的信元的TS/FID）按照特定插入/删除准则安置。图2给出了一种多层、多平面调度器200的实施例，它在图1所示ATM交换设备中实现调度功能，为MC流提供20 服务。各个聚合层的调度器200的每一个PQ实体都以树结构示出，得到每个业务平面的总嵌套树方案。下面的表给出了各个示例性业务以及相关的参数信息：

25

表 1

业务类别	应用	参数
实时恒定比特率 (CBR), 实时可变比特率 (rt-VBR)	语音(单信道或中继线), VBR 视频, 游戏	峰值信元率 (PCR), 信元时延变化容限 (CDVT)
非实时可变比特率 (nrt-VBR)	数据, 多媒体, 电邮, 视频流	PCR, CDVT, 可持续信元率 (SCR), 最大突发流量大小 (MBS)
保证帧频 (GFR)	附加数据, 浏览页面, 因特网	PCR, CDVT, MBS, 最大信元率 (MCR), 最大帧大小 (MFS)
尽力而为(未指定比特率或者 UBR)	廉价数据, 页面浏览和因特网	PCR, CDVT

5 与参数数据和 QoS 级别相关的其它细节可以参见分层调度器体系结构申请。继续图 2, 标号 206-1 到 206-6 分别指以下业务平面: rt[CBR/VBR] - 高平面, rt[CBR/VBR] - 中平面, rt[CBR/VBR] - 低平面, nrt-VBR 和 GFR 平面, GFR 平面, 以及 UBR (也就是, 可用最佳) 平面。标号 208 是指对应于子端口仲裁器 1305-i 的 PQ 树, 其中 PQ

10 树基于相应子端口所支持的 VC 连接生成。实际上, 调度器的每个子端口采用 (或者构造) 一个 PQ, 它是存储所有活跃 FID 的数据结构 (也就是与该 FID 相关的 VC 队列至少有一个信元)。所有子端口仲裁器的优胜者 (例如, 子端口仲裁器 222-i 和 222-j) 都被转发, 填充与总线级仲裁器 224-l 相关联的下一级 PQ 树结构 210。同样,

15 总线级仲裁器 224-l 和 224-k 转发各个选择给架级 PQ 结构 212。架

级仲裁器 214、栈面接口 216、传输层接口 218 转发它们的选择给管道级仲裁器 220，后者为特定业务平面选择一个优胜提名。

多种数据结构可以用于实现本专利申请所提出的层次调度中使用的树基 PQ。在本发明的一种示例性实施方式中，PQ 实体可以以堆结构实现。虽然堆实现通常在存储器使用方面表现优异，它受限于其算法复杂性，后者在高速设计中会使吞吐量受限。因此，在另一实现方式中，每一层特定的 PQ 实体作为综合“日历堆”结构实现，综合“日历堆”结构的详细描述由 Mudhafar Hassan-Ali 等人同一天提交的题为的“Calendar Heap System And Method For Efficient Sorting.”，美国申请号为 10/281,033；(代理卷号为: 1285-0101US)，该专利通过引用并入。

现在参看图 3，给出的高层框图说明了一种交换设备卡 300，它包括调度器块 310，用于按照本发明的教导实现接入节点中组播流调度方法。调度器块 310 包括 PQM 302 和 LBM 308，两者之间的多个接口实现与调度器操作相关的消息/数据通信。这些接口包括过期接口 314，优胜者流接口 316，堆/流重连接接口 318，以及堆/流插入接口 320。PQM 块 302 还与多个存储器块 304-1 到 304-4 接口，用以存储与上述分层调度器体系结构相关的各个 PQ 数据结构。连接到 PQM 块 302 的控制存储器 306 存储与 PQM 操作相关的控制程序代码。

LBM 块 308 也与一个或多个存储器块，例如存储器 312 接口，用以存储与 LBM 实现的监管和整形算法处理相关的信息。在一种实现中，LBM 用漏桶计算器作为状态机，它按照业务量合同和它所属连接的历史决定信元是否合格。在用于监管时，该状态机决定入信元是否符合要求，而用于整形时，它决定信元符合业务要求的时刻。根据业务类别的不同，提供了带有特定参数（也就是业务量描述符、漏桶参数（理论到达时间或 TAT、TS 值、信元到达时间等）的一个或多个特定算法处理，用以实现 LBM 的监管和整形操作。下面将会详细描述，本发明的 GFR 业务采用两种不同的算法 LB 处理来实现，每种算法具有不同的参数集，可分别用于调整保证流部分和非保证

流部分。

LBM 块 308 还与上下文存储器模块 (CMM) 332 和队列核心模块 (QCM) 332 接口, 用以实现其整体功能, 为该 ATM 交换设备服务的所有流维护漏桶信息。与 CMM 块 332 关联的信元到达接口 348 充当入信元的进入点。与 CMM 模块 332 相关联的上下文存储器 334 用于存储基于流的信息, 例如 QoS、FID、漏桶参数、目的路径标记 (DPT) 信息等。此外, 可以提供统计存储器块 336, 用以收集与该 ATM 交换设备卡 300 所服务的连接相关的性能监控数据。入口流上下文信息和出口流上下文信息分别通过接口 338 和 340 提供给 QCM 块 322。头/尾指针存储器 352 和统计存储器 354 连接到 QCM 块 322。与其相关联的信元指针接口 356 基于调度操作指向符合业务条件的信元。下面将针对组播流, 详细描述头/尾存储器和信元指针的功能。

QCM 块 322 和 LBM 块 308 之间的接口包括以下接口: 流激活接口 324、流重连接和去活接口 326、关闭连接接口 328 和超时接口 330。LBM 块 308 和 CMM 块 332 之间提供的优胜信元接口 342 用于发送优胜信元的相关信息。此外, 还提供了时钟管理块 344 和具有一个处理器接口 350 的处理器接口模块 346。

考虑到上述各种结构块, LBM 块 308 的总体功能包括以下各项:

流激活: 当信元到达或与其 FID 相关的信元缓冲器以前是空的时候, QCM 发送流激活消息给 LBM。该接口包含了与流相关的上下文存储器中发现的信息。

优胜流和流重连接/去活: 当 PQM 选择某个流作为优胜者时, 将该流从它所属的不同堆中删除 (例如根据聚合层)。如果特定流的信元缓冲中还有剩余的信元, 那么 LBM 重新计算 TS 值, 在 PQ 数据结构中重新连接该流。如果没有剩余的信元, LBM 计算该流的 TS 值, 将其存储在漏桶中, 直至该流的另一信元到达。

关闭连接: 如果 CAC 抑制某个流, 信元缓冲器会清除, FID 可以重用于其它连接。为了防止将旧的参数用于新连接, 必须通知 LBM 该流已不再有效。

超时：检查漏桶存储器，最好定期检查，以防止存储过期时间。

图 4 描述了一种示例性组播 (MC) 流方案，其中将进入 ATM 环境，例如图 1 所示接入节点 100 的 ATM 交换设备的根流，通过对应数量的叶子流发送给多个出口接口。在一种实施例中，可以使用调度器，例如前面提到的分层调度器体系结构申请中描述的调度器，来实现必需的 MC 调度功能。标号 402 代表了接收入口流 403 (也就是根) 的接口，入口流 403 需要调度，发送给 N 个出口接口 406-1 到 406-N。本领域中众所周知，入口流 403 被称为根流，发送给出口接口的 MC 流称为叶子流。标号 404 代表了与具有多个根流信元的根流 403 相关联的说明性队列。同样，标号 408-1 到 408-N 代表了与对应数量出口接口 (也就是叶子) 相关联的 N 个叶子流。下面会详细给出，每个叶子流需要初始化成指向指针存储器的索引指针，前述指针存储器实现了指针的循环链表，用以完成本发明的组播功能。例如，标号 410-1 到 410-N 代表了与 N 个叶子流相关联的 N 个叶子索引指针值。此外，每个叶子流带有一个 FID 和可应用的 LB 参数 (例如 TS)，当叶子流符合条件时，调度器利用它们发送复制的源业务量 (也就是根信元) 给相关的出口接口。

前面参照图 3 描述的 QCM 322 中，可以建立多达 256 个 MC 根，每个根最多可以有 2K 个叶子。一旦在 ATM 交换设备中建立了根流，就可以在 MC 会话中增删 MC 叶子。但是，在接受一个根流进入 QCM 之前，必须例如通过软件创建一个组播会话。在一种实施例中，MC 会话创建进程涉及利用属于各个 MC 会话的根流和叶子的上下文数据，对 CMM 332 相关联的上下文存储器进行编程。一旦对上下文存储器进行了编程，叶子可以在调度器模块 310 中激活。每个叶子可以由处理器利用 CPU 接口模块 346，通过在 QCM 块 322 中写入所需的寄存器来单独激活，如果不需要单点播送激活，QCM 块 322 就发送叶子激活消息给调度器模块 310。每个 MC 根也可以编程具有特定 MC 队列深度，它可以是 256、512 等。下面将看到，MC 队列深度与循环链表缓冲器结构中使用的指针数量有关，这些指针指向信

元存储缓冲器中用于存储根流信元的多个位置，顺序从根流的头信元开始。此外，MC 组的值和 MC 比特可以针对每个 MC 根设置，用以指示该流属于某个 MC 会话。需要理解，MC 叶子的编程需要将叶子与正确的根关联。在一种实施例中，这通过将 MC 叶子和同一会话的 MC 根编程为具有相同的 MC 组号来完成。此外，MC 比特和 MC 活跃比特设置为指示 MC 叶子对根活跃。MC 队列深度可以针对每个叶子编程设置，但一般来说这并不是必须的。如果没有为叶子设置 MC 活跃比特，调度器发送了一个非活跃叶子时，那么 QCM 为该叶子生成一个关闭连接命令。

10 在一种实施例中，QCM 进行以下操作，使用图 3 所示的功能块，将 MC 信元加入 ATM 交换设备的队列中，在上下文查找过程中，FID 用于检索流上下文，通过流上下文入口接口将它发送给 QCM。加入 MC 流所需的上下文数据是 MC 比特、MC 队列深度和 MC 组号。CMM 最好能够根据流信元到达计数来更新。流上下文入口接口内的 MC 比特使得 QCM 能够识别属于 MC 流的流。QCM 产生一个地址，指示了头/尾内正确的 MC 表和指针存储器，下面将给出详细描述。头/尾存储器的索引可以用于检索根序列比特和当前根流索引，它是根写索引指针。开始时，整个头/尾存储器都可以设置归 0，从而使得所有 MC 会话的当前根索引值为 0。在一种实现中，指针存储器的 MC 地址可以根据 MC 组号、当前根索引以及偏移值生成。这样，可以看出，指针存储器是以生成的 MC 指针存储器地址为索引，其中在验证了位置有效性之后，将根信元指针写入指针存储器。

20 为了能够从 ATM 交换设备的队列中去除组播信元，QCM 块执行以下操作。调度器模块发出通过优胜信元接口 342 发送的流。利用优胜流的 FID 作为地址，检索得到上下文，通过出口流上下文接口 340 发送给 QCM 块。从队列中去除一个 MC 信元所需的上下文数据是 MC 比特、MC 活跃比特、叶子序列比特、MC 队列深度以及 MC 组号。操作 CMM 块以更新流信元发送计数，因为信元在优胜仲裁之后就进行发送调度。该接口中的 MC 比特使得 QCM 块能够识别该

流属于 MC 流。如果 MC 活跃比特=0，则向调度器发送一个关闭连接命令；否则，QCM 块产生一个地址，指示了头/尾内当前 MC 表和指针存储器。QCM 利用生成的地址，指向头/尾存储器，并检索根序列比特和当前根索引。该信息与出口流上下文数据一起用于将信元从 MC 队列中去除。QCM 块随后产生一个指示了适当的 MC 表的地址，并发送叶子索引指针所指示的信元的副本，前述叶子索引指针是一个读指针。类似于根信元入队列处理，该指针存储器的 MC 地址由 MC 组号、叶子索引值以及偏移值生成。

在发送信元副本时，如下所述更新叶子索引值和叶子序列。当叶子索引等于 MC 队列深度时，叶子到达了 MC 队列尾端，因此需要翻转。如果叶子索引等于当前根索引，叶子读指针等于根信元写指针，不为该叶子发送根信元副本；而是从交换设备发送一个空闲信元。另一方面，如果叶子索引不等于当前根索引，叶子读指针不等于根信元写指针，没有失序问题，从交换设备发送该指针（向其引用叶子索引）指向的信元副本，叶子索引增加 1。

当叶子和根失序时，发送当前根索引的信元副本，如下所述更新叶子索引。如果当前根索引等于 MC 队列深度，叶子索引就初始化为 0。另一方面，如果当前根索引不等于 MC 队列深度，叶子索引等于当前根索引增加 1。

在一种实现中，在发送信元副本时，指针不返回空闲指针表，前述信元副本位于 MC 指针所指向的信元存储器位置。叶子索引值和叶序列比特都在上下文存储器中根据叶子 FID 更新。

现在参看图 5，该图给出了实现本发明 MC 队列缓冲器系统的一种存储器实施例 500。存储器实施例 500 包括头/尾存储器 502，指针存储器 504 和信元存储器 506，用于实现前面详细描述 MC 调度操作。如前面给出，叶子索引值 410-1 到 410-N 对应于与根流 403 相关联的 N 个叶子流，它们在头/尾存储器 502 中初始化。指针存储器 504 包括具有多个指针 512-1 到 512-M 的循环链接指针缓冲器 510，用于支持深度为 M 的 MC 队列。叶子索引指针位置包含的索引值称

为链接 MC 指针，因此，它作为叶子读指针，这在前面解释过。因此，叶子索引值可以认为是索引到各个叶子流的信元位置的“头”，在优胜仲裁之后，它们从这些位置得到信元副本。同样，根索引指向的指针位置作为根信元写指针，可以认为是 MC 队列的“尾”，当信元到达 ATM 交换上下文时，这些信元会被加入前述 MC 队列。

指针 512-1 到 512-M 可以以任何目前为止已知或未知的技术循环链接。链接缓冲器 510 包括与叶子和根流索引值相关的索引部分 514A，以及指针位置部分 514B，它包含到信元存储器位置 516-1 到 516-K 的指针，它们可以是连续的，分散的，或者通过其他方式进行组织。根信元利用根索引存储在，或写入信元存储器位置，处理从头根信元和附加根流信元开始，因为它们已在 MC 业务的队列中。当特定叶子流符合调度条件时（例如基于其 LB 参数，例如 TS），得到头根信元的副本进行发送，并更新优胜流的叶子索引。当该流再次赢得仲裁时，针对信元存储器中的下一信元副本进行发送调度。

图 6 给出了一种循环链接指针缓冲器系统 600 的实施例，用于为具有 8 个存储器信元位置深度的 MC 队列提供服务，每一个存储器信元位置都由一个链接指针指着。最初，叶子流索引值和根流索引都是同一个公共指针。在头根信元到达时，根流索引所包含的指针指向可以存储头根信元的信元存储器位置。更新根流索引，使得在附加根信元到达时，它们能存储在适当的信元位置。在存储了 8 个根信元之后，根流索引可以停止在当前根流索引。叶子流在基于各自的 LB 参数作出优胜仲裁之后，开始根据各个索引值得到信元副本，各个索引值也在循环指针缓冲器中更新。因为针对各个时隙独立进行仲裁，叶子流可以独立得到根信元的副本，而不会相互干扰。这样，一旦 MC 会话开始，叶子流索引值可以以不同速率推进，指向不同的循环指针。例如，叶子流 ID2 和叶子流 ID3 的索引都是指向指针 1，叶子流 ID1、叶子流 ID5、叶子流 ID4 和叶子流 ID6 分别指向指针 2、指针 3、指针 4 和指针 6。

图 7 给出了本发明组播调度方法的一种实施例中采用的操作流

程图。在框 702、704 和 706 中，执行在 ATM 环境中，例如接入节点的 ATM 交换设备中开始 MC 会话的不同的初始化过程。本领域中的技术人员应当理解，这些初始化过程可以由具有任何顺序的若干步骤实现。因此，下面有关初始化过程的描述应当被认为只是说明性的。在通过节点的交换设备接收根流信元进行调度时，相应于与 MC 会话相关的出口接口，初始化多个叶子流 ID (框 702)。初始化循环链接的 MC 队列 (MCQ) 指针列表，每个链接指针指向一个信元存储器位置。利用根流索引，将根流信元写入信元存储器位置，处理从头根信元开始 (框 704)。叶子流 ID 的索引是指向头根信元位置的 MCQ 指针，从而在根流的正确开始处，开始信元的发送 (框 706)。

在特定叶子流变得符合调度条件时，得到头信元内容的副本，发送给与优胜叶子流相关联的出口接口。更新叶子流的索引指针，使其指向 MCQ 的下一信元的 MCQ 指针 (框 708)。一旦为所有叶子流调度了特定信元 (也就是说，每个出口接口都得到了特定信元的副本)，该信元的信元存储器位置可以利用更新 (也就是当前) 根流索引，写入新的根信元 (框 710)。这些操作可以继续，直至已针对每个叶子流通过该交换设备发送所有的根流信元。

基于前面的讨论，应当理解，本发明提供了在 ATM 环境中实现组播方法的一种创新方案，前述 ATM 环境可以是接入节点中的交换设备，也可以是区域范围内的传输网，其中有效地克服了当前实现的不足和缺陷。通过提供单个 MC 信元缓冲器，该缓冲器可以被多个叶子流独立访问，从而不再需要支持多个队列。此外，这种独立方案支持分离流，从而消除了流间干扰。另外，本领域技术人员很容易理解，尽管针对接入节点的 ATM 交换设备的分层调度器提出了本发明的组播方案，但这里包含的教导并不局限于该环境；它们也可以在其它 ATM 应用中实现。

通过前面的详细描述，相信本发明的操作和构造已经很清楚。示出和描述的本发明实施例是例示性的，应当容易理解在不偏离后附

权利要求书所提出的本发明范围的前提下，可以作出各种变化和
改进。

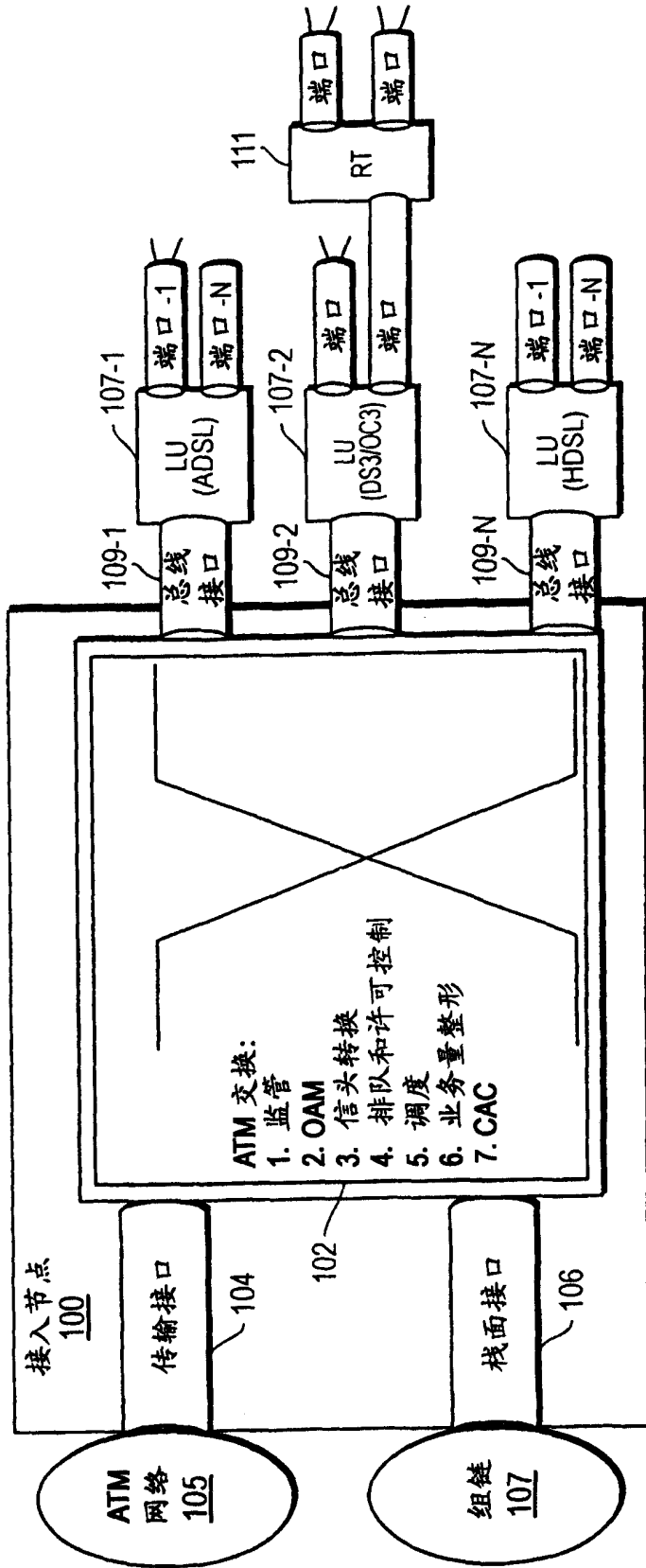


图 1

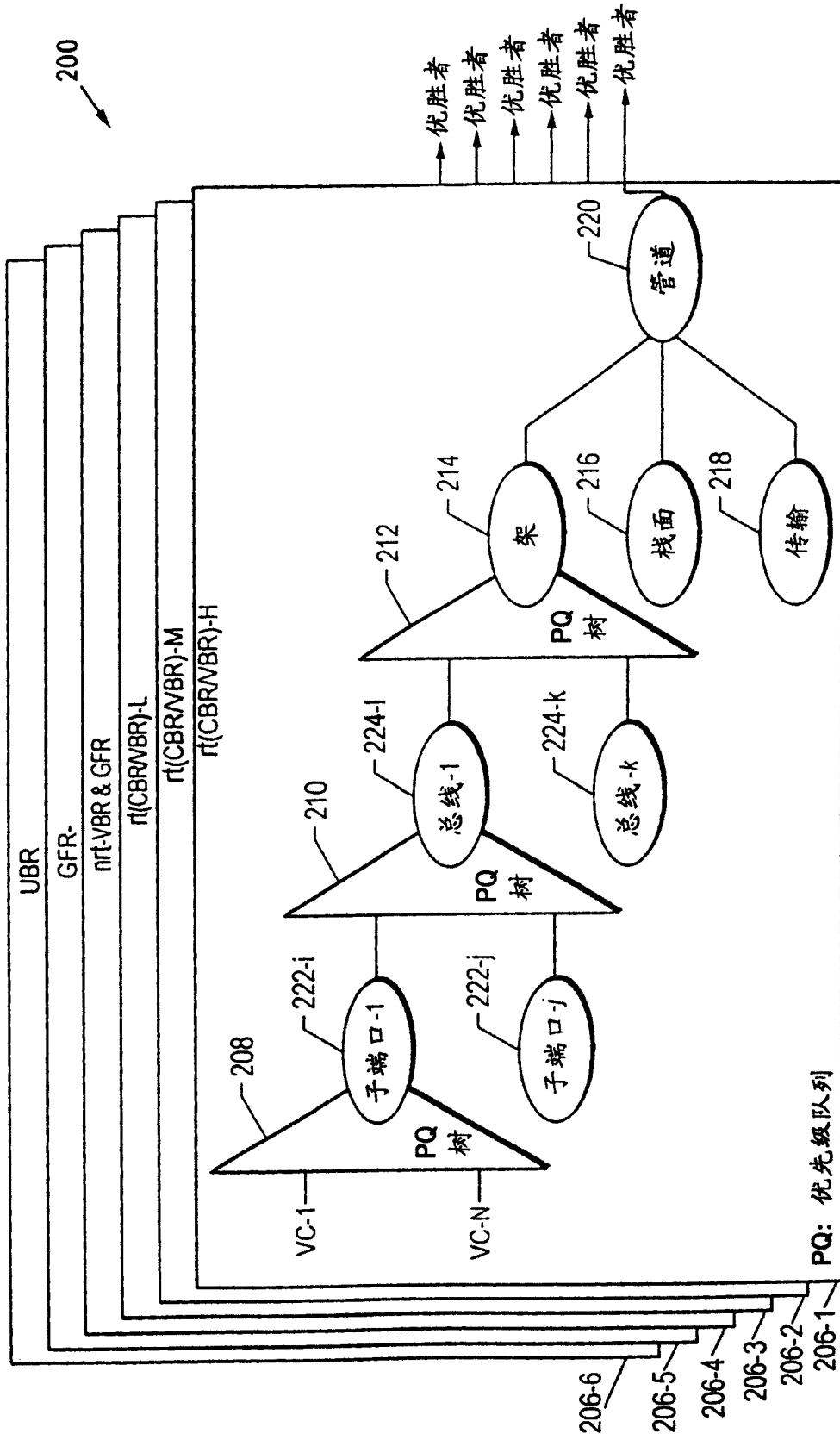


图 2

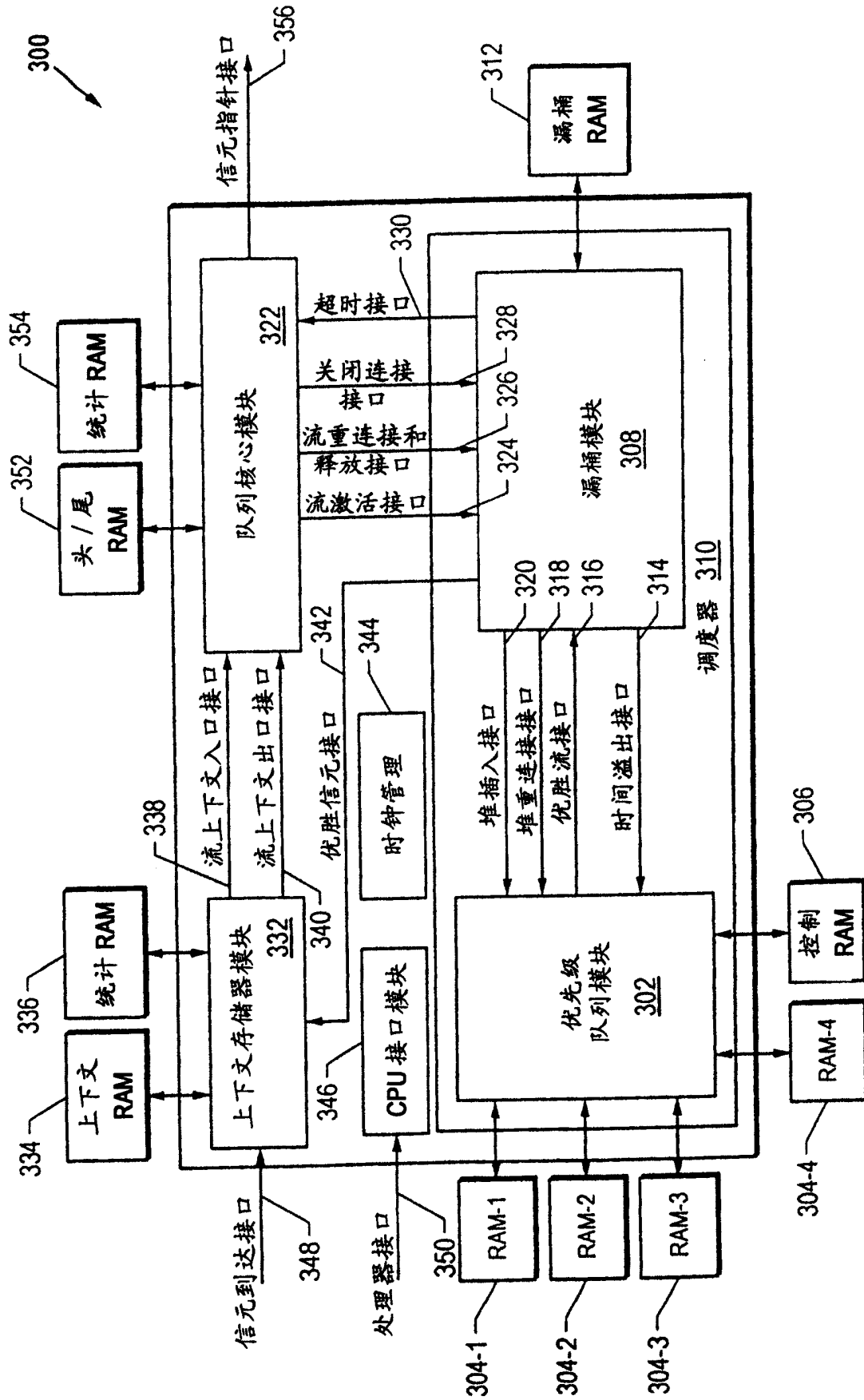


图 3

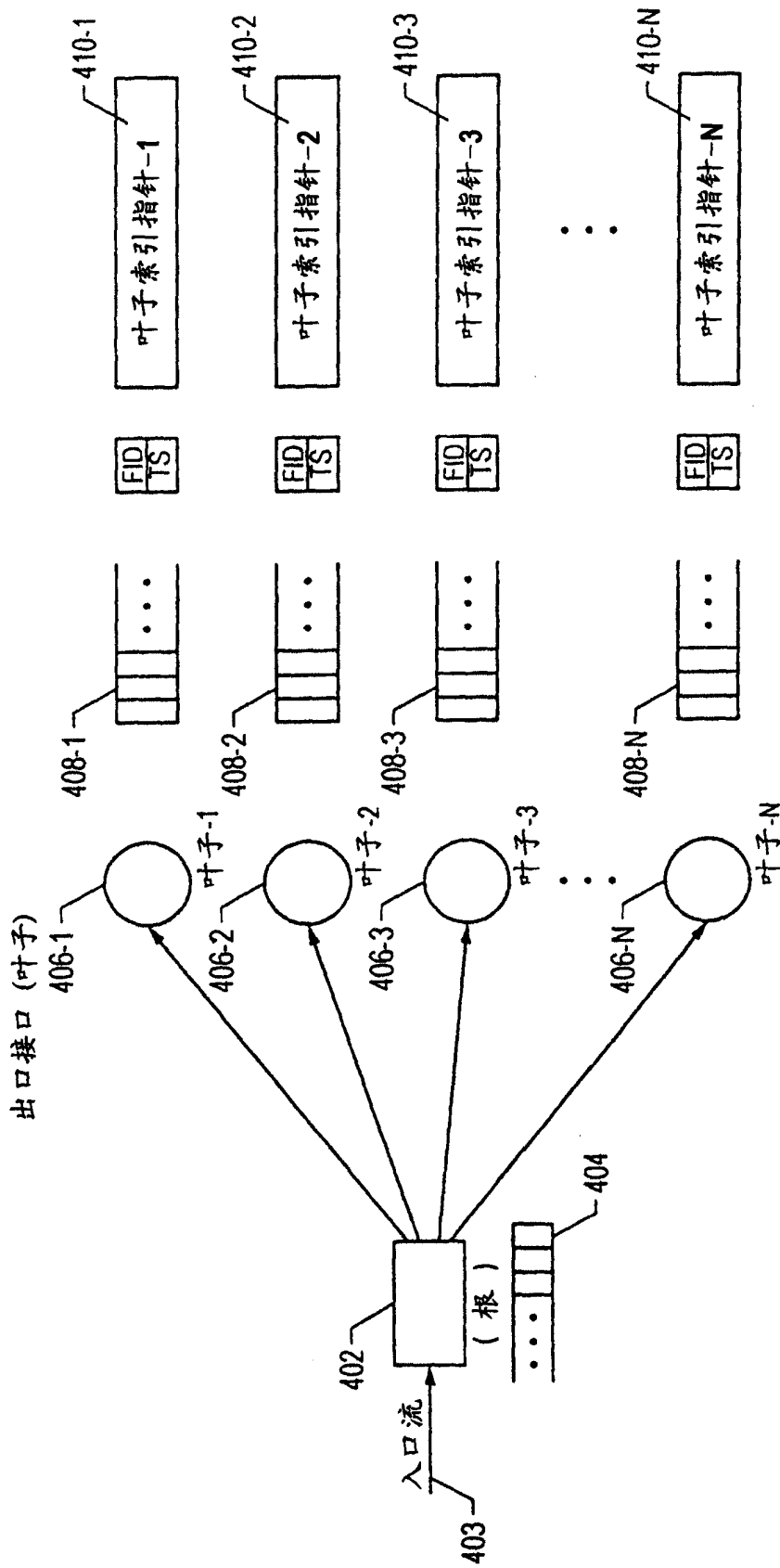


图 4

500 ↗

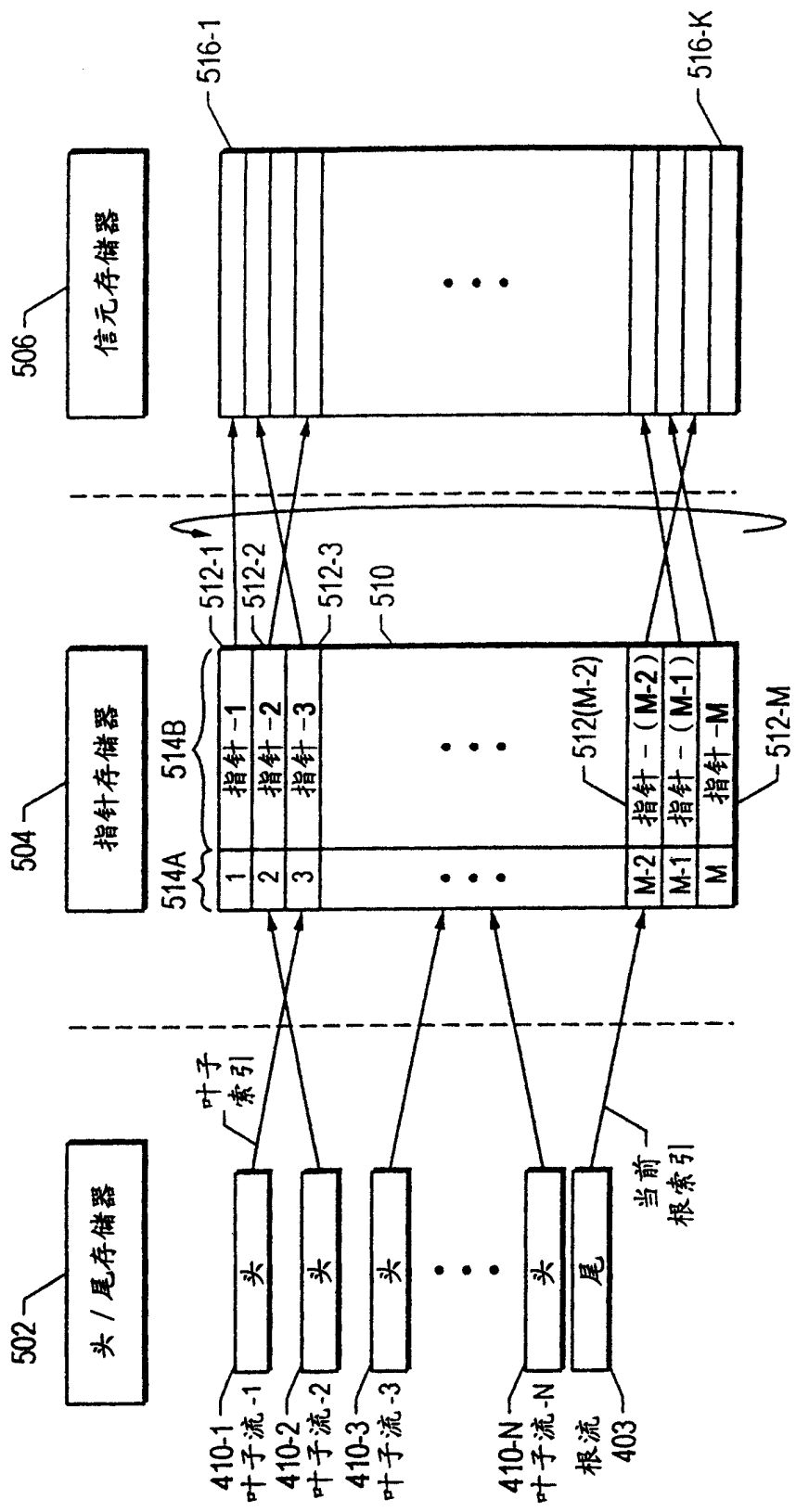


图 5

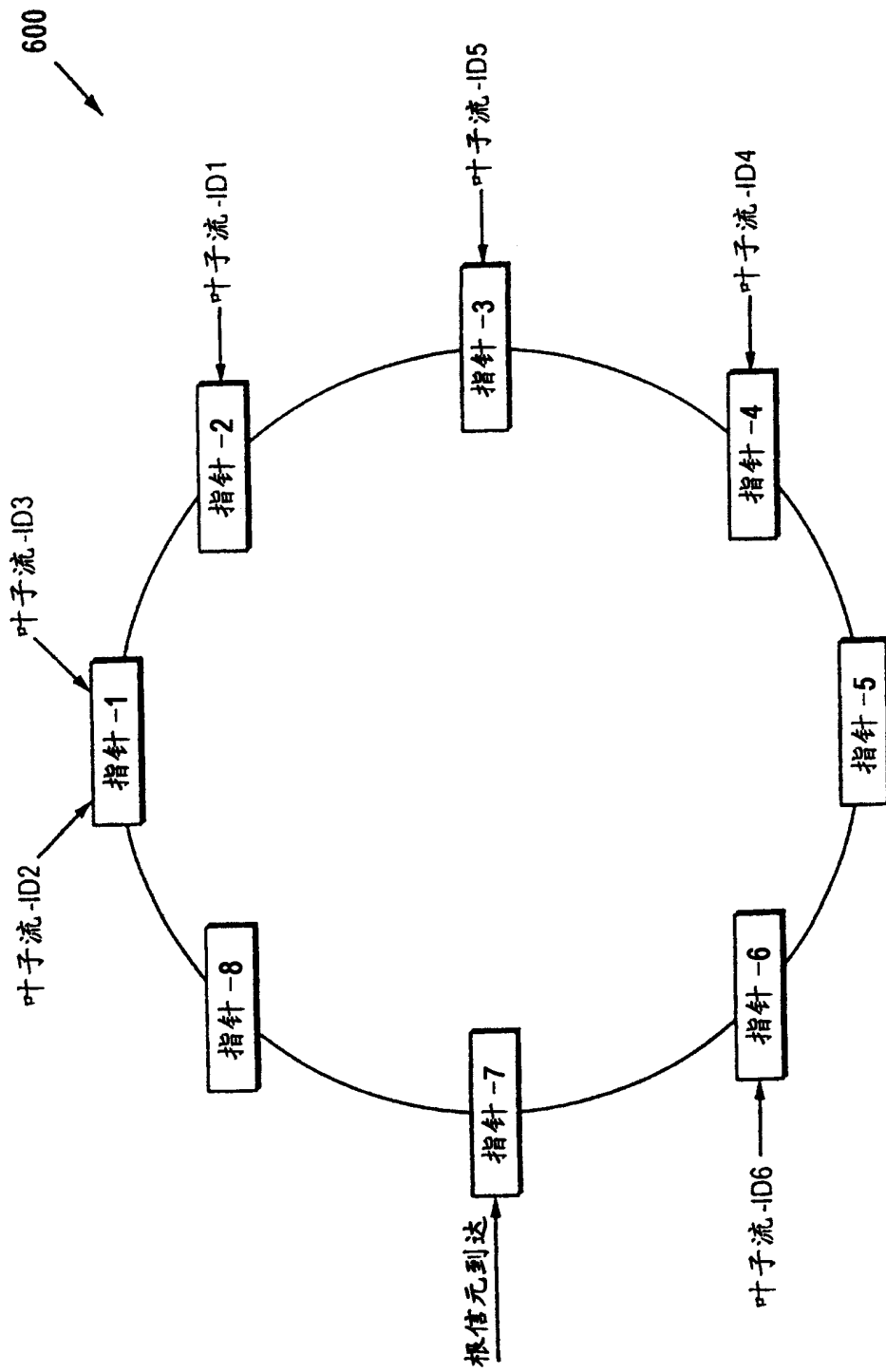


图 6

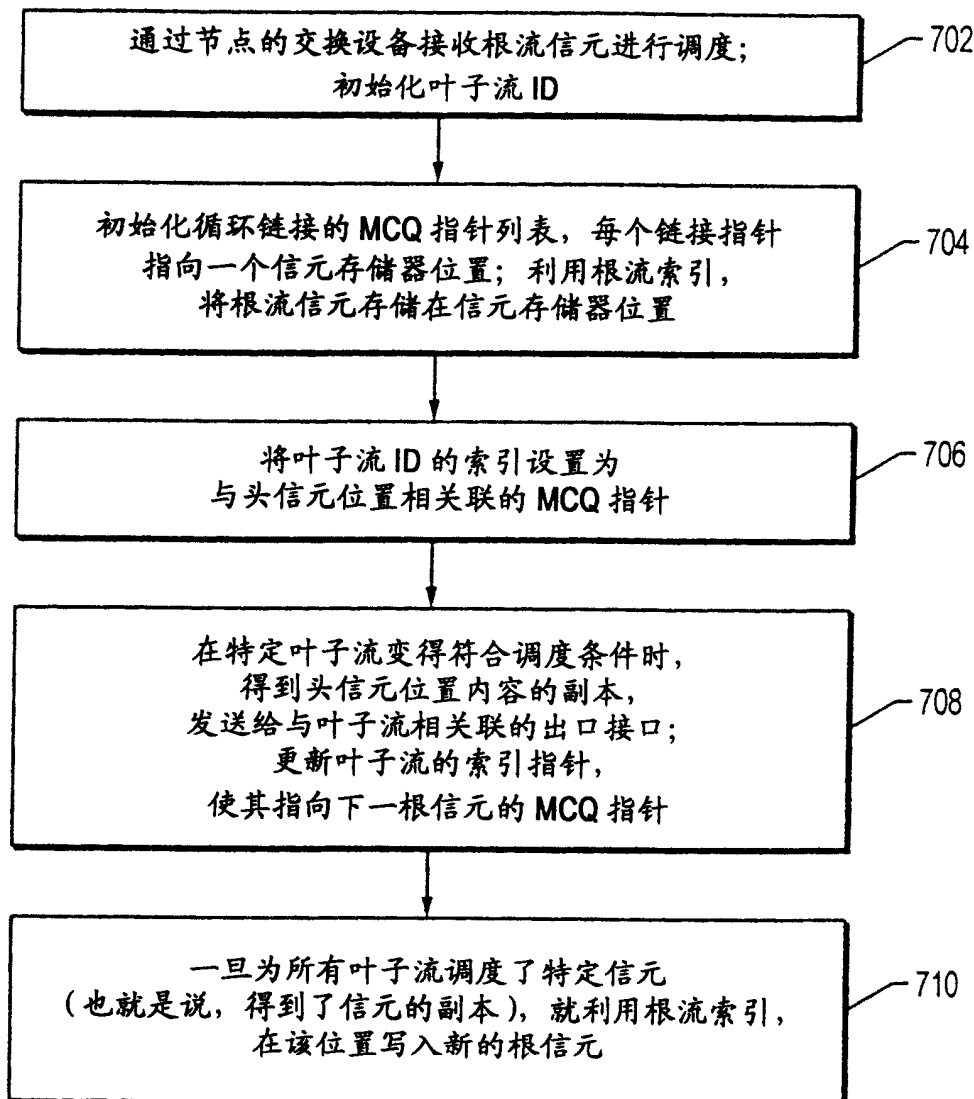


图 7