

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H04Q 11/04

H04L 12/56 H04L 12/64



[12] 发 明 专 利 说 明 书

[21] ZL 专利号 99801400.1

[45] 授权公告日 2004 年 9 月 8 日

[11] 授权公告号 CN 1166247C

[22] 申请日 1999. 6. 18 [21] 申请号 99801400. 1

[30] 优先权

[32] 1998. 6. 19 [33] US [31] 60/090,028

[86] 国际申请 PCT/US1999/013641 1999. 6. 18

[87] 国际公布 WO1999/066758 英 1999. 12. 23

[85] 进入国家阶段日期 2000. 4. 19

[71] 专利权人 杜松网络公司

地址 美国加利福尼亚州

[72] 发明人 F·卡斯滕霍兹 T·维斯特伯格

S·R·维利斯

审查员 李振华

[74] 专利代理机构 中国专利代理(香港)有限公司

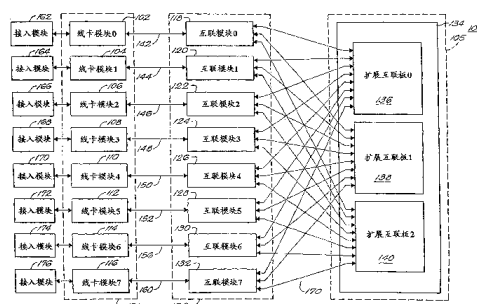
代理人 吴立明 王忠忠

权利要求书 6 页 说明书 30 页 附图 12 页

[54] 发明名称 通讯节点、通讯互联网络和在其中
传输信号的方法

[57] 摘要

一种在通信节点中运行的互联网络，其中互联网络有传送不同通信协议，带宽可调和减低关机时间的能力。根据本发明的一个实施例，该通信节点包括多个 I/O 信道，用于耦合信息进出该节点，和该互联网络包括至少一个本地互联，带有本地传送部件而在多个 I/O 信道之间传送信息；和调节部件，用于扩展互联以包括附加互联模块，使得信息能够在包含在该互联网络中的本地互联模块间传送。



1. 一种用于在通讯节点内工作的互连网络,所述网络包括:

多个本地线卡模块,它们被配置成处理以多种速度所接收且依据多种协议格式化的信息,

数目可选的本地互联模块,它们被连接到所述本地线卡模块且在位置上彼此靠近,所述本地互联模块中的每一个都包括本地传输单元,用于在多个本地 I/O 信道之间传输信息和在所述多个本地 I/O 信道和多个非本地 I/O 信道之间传输信息,以及

扩展互联模块,在位置上靠近所述本地互联模块,并且包括耦合装置,用于与所述非本地 I/O 信道进行电耦合,和扩展传输单元,用于在所述本地互联模块之间传输信息。

2. 根据权利要求 1 所说的互连网络,其中,所述本地传输单元包含信息同步装置,用于同步在所述本地传输单元中的每个本地传输单元之间所传输的信息。

3. 根据权利要求 1 所说的互连网络,其中,所述扩展传输单元包含信息同步装置,用于同步在所述本地互联模块之间所传输的信息。

4. 根据权利要求 1 所说的互连网络,还包含热插拔装置,用于在所述互连网络正传输着信息的同时,更换在所述互连网络中包含的数目选定的所述本地互联模块。

5. 根据权利要求 1 所说的互连网络,其中,所述本地 I/O 信道具有相关联的优先级,并且所述互连网络进一步地包含质量服务装置,用于传输来自所述本地 I/O 信道中具有相对较高优先级的一个本地 I/O 信道的信息,上述的信息传输要比传输来自所述本地 I/O 信道中具有相对较低优先级的一个本地 I/O 信道的信息优先进行。

6. 根据权利要求 1 所说的互连网络,其中,所述本地 I/O 信道具有用于接收信息的相关可用性,并且所述互连网络进一步地包含状态装置,用于保持表示一个或多个所述本地 I/O 信道的所

述相关可用性的状态，和反压装置，用于将所述状态从所述本地 I/O 信道中传送出去。

7. 根据权利要求 6 所说的互联网络，其中，所述本地互联模块进一步地包含多个内存队列，所述多个本地 I/O 信道具有相关的内存队列，并且所述内存队列包含存储装置，用于存储通过相关 I/O 本地信道所接收的信息。

8. 根据权利要求 7 所说的互联网络，进一步地包含设置装置，用于响应一特定本地 I/O 信道的相关队列到达可选择的容量级别，设置与所述特定本地 I/O 信道相对应的状态，以表示用于接收信息的不可用性。

9. 根据权利要求 7 所说的互联网络，其中，被耦合进和耦合出所述网络的信息具有相关的优先级，并且所述网络进一步地包含设置装置，用于响应所述相关队列到达可选择的容量级别，设置与一个特定的所述本地 I/O 信道相对应的状态，以表示用于接收具有一特定优先级的信息的不可用性。

10. 根据权利要求 6 所说的互联网络，其中通过所述互联模块传输的信息包含一个数据部分和一个目的地址部分，所述目的地址部分指定要通过所述本地 I/O 信道中的哪一个本地 I/O 信道来将所述信息从所述互联网络中传送出去，并且所述反压装置适用于将所述目的地址部分替换成所述状态，以便将所述状态从所述互联网络中传送出去。

11. 根据权利要求 1 所说的互联网络，进一步地包含冗余生成装置，用于生成正被通过所述本地 I/O 信道向所述互联网络之外传输着的信息的另一版本。

12. 根据权利要求 11 所说的互联网络，其中，所述互联网络适用于以信息信元来传输信息，所述信元中的每一个信元包含信息字的组，并且所述冗余生成装置适用于使得所述信息的所述另一版本成为在一个信息信元中所包括的所述信息字组对之间逐位的“异或”。

13. 根据权利要求 1 所说的互联网络，其中，所述互联网络适用于以信息信元来传输信息，并且所述本地传输单元进一步地包

含传输装置，用于基本上同时地将多个所述信息信元从所述本地互联模块中传送出去。

14. 根据权利要求 13 所说的互联网络，进一步地包含内存队列，用于在中间存储要加以传送的所述多个信息信元，和队列检测装置，用于检测选定数目的所述信息信元被存储到所述内存队列之中的时间，其中，所述本地传输单元适用于响应所述队列检测装置检测到所述选定数目的信息信元被存储而传输所述多个信息信元。

15. 根据权利要求 1 所说的互联网络，其中，所述互联网络适用于以信息信元来传输信息，并且所述扩展传输单元进一步包含传输装置，用于基本上同时地将多个所述信息信元从所述扩展互联模块中传送出去。

16. 根据权利要求 15 所说的互联网络，进一步地包含内存队列，用于在中间存储要加以传送的所述多个信息信元，和队列检测装置，用于检测选定数目的所述信息信元被存储到所述内存队列中的时间，其中，所述扩展传输单元适用于响应所述队列检测装置检测到所述选定数目的信息信元被存储而传输所述多个信息信元。

17. 根据权利要求 1 所说的互联网络，其中，所述本地 I/O 信道具有相关的目的地址，并且所述扩展互联模块包含：

数组装置，用于存储表示通过所述扩展传输单元的多个路径的路径信息，通过其能够将信息从所述本地互联模块中的一个本地互联模块的第一本地 I/O 信道传输到所述本地互联模块的另一个本地互联模块的第二个本地 I/O 信道，和

索引装置，用于至少部分地响应所述第二个本地 I/O 信道的目的地址，从所述数组装置中选择合适的路径信息。

18. 根据权利要求 1 所说的互联网络，其中，所述本地传输单元和所述扩展传输单元基本上相同。

19. 根据权利要求 18 所说的互联网络，其中，所述本地传输单元和所述扩展传输单元都包含模式控制装置，用来选择将所述传输单元应用到所述本地互联模块中的一个本地互联模块中还是

应用到所述扩展互联模块中。

20. 一种可动态扩展的通讯互连网络，包含：

数目可选的本地互联模块，每一个本地互联模块都包括相关的传输单元，用来将信息传输通过所述相关的本地互联模块，以及

单一的扩展互联模块，包含耦合单元，用于耦合所述数目选定的本地互联模块，和扩展传输单元，用于在所述本地互联模块之间传输信息，其中，所述数目可定的本地互联模块能够在所述扩展互联模块正传输着信息的同时加以变更。

21. 一种用于在通讯节点内工作的互连网络，包含：

数目可选的本地互联模块，每一个本地互联模块都包括本地传输单元，用于在多个本地 I/O 信道之间传输信息和在所述多个本地 I/O 信道和多个非本地 I/O 信道之间传输信息，其中，所述互连网络适用于以信息信元来传输信息，所述信元中的每个信元都包括信息字的组，

扩展互联模块，包括耦合装置，用于与所述非本地 I/O 信道进行电耦合，和扩展传输单元，用于在所述本地互联模块之间传输信息，以及

冗余生成装置，用于通过生成一个信息信元中所包括的所述信息字组对逐位的“异或”来生成从所述互连网络向外传送中的信息的另一版本。

22. 一种用于在通讯节点内工作的互连网络，包含：

数目可选的本地互联模块，每一个本地互联模块都包括本地传输单元，用于在多个本地 I/O 信道之间传输信息和在所述多个本地 I/O 信道和多个非本地 I/O 信道之间传输信息，

扩展互联模块，包括耦合装置，用于与所述非本地 I/O 信道进行电耦合，和扩展传输单元，用于在所述本地互联模块之间传输信息，以及

服务质量装置，用于根据所述信息的相关优先级区分被耦合到所述本地 I/O 信道中的信息，和确定用于接收在一个或多个所述本地 I/O 信道上具有一特定相关优先级的信息的不可用性。

23. 一种用于在通讯节点内工作的互联网络, 所述网络包含:

数目可选的本地互联模块, 每一个本地互联模块都包括本地传输单元, 用于在多个本地 I/O 信道之间传输信息和在所述多个本地 I/O 信道和多个非本地 I/O 信道之间传输信息, 以及

至少一个扩展互联模块, 包括耦合装置, 用于与所述非本地 I/O 信道进行电耦合, 和扩展传输单元, 用于在所述本地互联模块之间传输信息,

其中, 所述互联网络适用于以信息信元来传输信息, 并且所述本地和扩展传输单元进一步包含丛装置, 用于基本上同时地传输多个所述信息信元。

24. 一种带宽可动态扩展的通讯节点, 包含:

数目可选的本地通讯模块, 所述本地通讯模块给所述通讯节点提供带宽的动态可扩展性, 所述本地通讯模块中的每一个通讯模块都包括用于耦合进出所述节点的信息的多个外部通讯信道, 用于在所述节点之内耦合信息的多个内部通讯信道, 以及用于在所述外部通讯信道和所述内部通讯信道之间耦合信息的装置,

本地互联模块, 与所述数目可选的本地通讯模块中的每一个本地通讯模块相关, 并包括本地传输单元, 用于在所述相关的本地通讯模块的所述多个内部通讯信道之间传输信息, 和

扩展互联模块, 包含耦合装置, 用于与每一个所述本地通讯模块相耦合, 和扩展传输单元, 用于在所述本地通讯模块之间传输信息, 由此就能够在所述数目可选的本地通讯模块中的不同本地通讯模块的外部通讯信道之间传输信息。

25. 一种带宽可动态扩展的通讯节点, 具有模块化的结构并且包含:

数目可选的本地通讯模块, 所述本地通讯模块给所述通信节点提供带宽的动态可扩展性, 所述本地通讯模块中的每一个本地通讯模块都包括多个通讯信道, 用于将信息传送到所述节点中和将信息从所述节点中传送出去,

本地互联模块, 与所述本地通讯模块中的每一个本地通讯模块相耦合, 并包括本地传输单元, 用于在所述多个通讯信道之间传

输信息，以及

扩展互联模块，包含耦合单元，用于与多个所述本地通讯模块相耦合，和传输单元，用于在所述多个本地互联模块之间传输信息，其中，信息可以在所述本地通讯模块的不同通讯模块的通讯信道之间传输。

26. 一种在适用于在通讯节点内工作的互联网络内的 I/O 信道间传输信息信号的方法，包含以下步骤：

将信息耦合到一个或多个位置上靠近的本地互联模块的 I/O 信道上，

响应本地目的地址，在所述本地互联模块中的一个特定本地互联模块的 I/O 信道间传输信息，

响应非本地目的地址，将信息从所述本地互联模块中的一个本地互联模块传输到位置上靠近的扩展互联模块的 I/O 信道，

响应所述非本地目的地址，将信息从所述扩展互联模块传输到所述本地互联模块的另一个特定的本地互联模块，以及

通过在所述多个本地互联模块中包含数目可选的所述本地互联模块来调节所述互联网络的带宽。

通讯节点、通讯互联网络和在其中传输信号的方法

5 技术领域

本发明一般涉及通讯节点，特别涉及可在通讯节点内工作的互联网络。

背景技术

10 作为在多个信号源之间传输通讯信号的结合点，通讯节点需要处理多个流行的通讯协议，比如综合业务数字网（ISDN）协议，异步传输模式（ATM），以及互联网协议（IP）。ISDN 作为一种早期的对多服务架构（即一种可处理多种通讯信息封装的架构）的尝试，是基于电话系统的，在线路上被分配 64 千比特每秒（Kbps）的带宽。相对于在工作带宽为 10 兆比特每秒（Mbps）的局域网（LANs），ISDN
15 被证实是太慢了。被认为是适用于宽带 ISDN 传输机制的 ATM 是一种包交换协议。ATM 传输信息用的是被称为信元的固定长度的包。信元在通讯节点之间的在每一个通讯会话建立之前就建立好的虚拟连接（VC）上传输。固定信元格式和虚拟连接的结合使得 ATM 成为 ISDN 的高速替代者。此外，ATM 处理突发数据流量的效率比时分多路（TDM）更高，并且还可以提供高质量的语音和视频支持。
20

万维网（WWW）的流行促进了 IP 的使用。结果，价格低廉，不受距离限制的基于 IP 的传输方式成为了租用线路与帧中继（FR）的具有竞争力的替代者。此外，互联网服务提供商（ISP）成为了企业数据传输的有力竞争者。

25 不幸的是，以上所有的信息传输协议都有缺点。特别是，ISDN 提供了一种比较慢速的接入方案。ATM 支持帧中继，虚拟私有网（VPN），线路仿真，专用分组交换机（PBX）的互联和服务质量控制（QoS），但并不能使现存的各种数据协议很方便地结合起来。IP 支持诸如互联网接入和 VPN 之类的认为连接性很重要的应用。不过，
30 IP 还是需要证明它具有工业级的可靠性。结果，全面服务提供商发现维持并行的交换网络是必要的。因为这样的并行网络需要对诸如语音交换机，帧中继交换机，ATM 交换机，路由器，加/减多路复用

器和数字交叉连接器之类的各种设备提供维护和服务，它们具有很高的相关的固定资产费用。

5 常规的通讯节点也有多个缺点。例如，常规的通讯节点不能很方便地提供可扩展性。一般来说，常规的交换节点提供了一种具有固定带宽的交换/路由网络。可是，随着企业的发展，它的需求也在增长。但是，基于常规技术的固定带宽的交换路由网络需要企业能够事先预知这样的增长，并且能事先购买具有足够带宽的系统；这就带来了维持并行的网络这一挑战。常规系统的另外一个缺点在于可靠性。由于常规系统没有提供一种可以在各种协议上都能运行的单一交换/路由网络，今天的千兆出现点（Giga Points-of-Presence）
10 （GigaPoPs）和 Access PoPs 成了连接小型 Access PoPs 与核心传输容量相连的核心路由器的复杂且昂贵的集合。这种结构是脆弱的，性能的限制和设备的失灵将导致频繁的服务故障。由于故障或与常规技术相关的升级所造成的可观的停机时间是企业所不能承受的。

15 由于常规系统的交换/路由网络主要是为在特定协议限制下工作所设计，它们缺乏为适应新生技术以及使用新的通讯协议所需的灵活性。正如前面所讨论的，不同的协议提供了不同的 QoS 特性。这样，在单一协议限制下工作的网络的另一个缺点是，对于具有不同优先级需求的用户服务提供商不能提供不同级别的服务；这使得服务提供商不得不事先放弃一个潜在的且重要的收入来源。

20 发明内容

此项发明的一个目的就是要提供一种能够在多服务通讯节点上处理不同通讯协议的互联网络，而不需要维持那些昂贵的并行的网络。

25 此项发明的一个进一步的目的是要提供一种能使通讯节点适应采用新生技术的通讯协议的互联网络。

此项发明的另一个目的是要提供一种能根据供应商不同的带宽需求调节通讯节点的带宽的可扩展的互联网络。

此项发明的另一个目的是要提供一种能够在修复与升级时，不会造成停机或是通讯节点工作性能下降的容错的互联网络。

30 此项发明的这些或是其他目的将会在下面的关于本发明的描述中变得显而易见。

本发明涉及通讯节点。更确切地说，涉及通讯节点中的互联网络。

根据本发明的一种实施方案，一个通讯节点包含能使这个节点传输不同通讯协议的互联网络。根据更进一步的实施方案，根据本发明的互联网络能够使通讯节点用相同的硬件处理 ATM 和 IP Packet-over-SONET 协议。根据本发明的另外的实施方案的，互联网络能够使
5 使通讯节点提供帧中继数据终端设备（DTE）和多协议标志转换（MPLS）的功能。根据本发明的另一种实施方案，互联网络能够使通讯节点成为既是本地 ATM 交换机又是本地 IP 路由器，其工作线速度起码能达到 2.488Gps（OC48c/SM16c）。

根据附属的特性，本发明能够进一步的提高可靠性。举例来说，
10 根据一种实施方案，本发明提供自动保护交换（APS），在开放式系统互联（OSI）的第 2 层和第 3 层信息将被映像以提供快速的 APS 转换。此外，系统模块将是可热插拔的，并被设计成当单一组件发生故障时不会导致整个节点瘫痪。

根据另一种实施方案，通讯节点将被包装成模块的可扩展集合。
15 OC48 线卡和千兆位以太模块将包含一个本地互联网络而成为本地通讯接口模块。一个可选的前端接入模块将提供诸如 OC12/STM4, OC3/STM1, DS3, 或是 E3 的接口，一个可选的扩展互联模块，有时也被称作超连接结构(hyperconnect fabric)，将允许通讯节点的带宽动态扩展到多达八个互联的本地互联模块，这样就可以提供基本上不会阻塞的 160Gbs 的带宽。
20

本发明的另外一种实施方案能够使服务商提供为企业提供不同程度或质量服务（QoS）。

简要地说，根据本发明的一种实施方案的一个互联网络与一个具有一个本地通讯接口，一个相关本地互联网络和可扩展单元的通讯
25 节点合为一体。本地通讯接口包含多个耦合进出节点信息的外部通讯信道和多个在节点内部传输信息的内部通讯信道。每一外部通讯信道都与一内部通讯信道耦合。本地互联网络有在内部通讯信道之间传输信息的本地传输单元，结果信息得以在外部通讯信道之间传输。可扩展单元可以动态地扩展节点使其可以包含具有辅助相关本地
30 互联网络的辅助本地通讯接口，这样信息就可以在本地通讯接口之间传输了。根据一个进一步的特性，在一个节点扩展以包含辅助本地通讯接口和本地互联网络以后，此通讯节点就可以可选地在任

何的内部通讯信道间传输信息，因而对于本地通讯接口任何的外部通讯信道和也是如此。

5 由于本发明的一个增强特性是动态的带宽可扩展性，根据进一步的实施方案，通讯节点提供辅助本地通讯接口和辅助本地互
网络；可扩展单元包含一个扩展的互网络。辅助本地通讯接口包含
多个耦合进出节点信息的辅助外部通讯信道和多个在节点内部传输
信息的内部通讯信道。辅助本地通讯接口的这些内部和外部通讯信
道互相耦合在一起。辅助本地互网络包含为在多个辅助内部通讯
信道间传输信息而设置的辅助本地传输单元。本地互网络和辅助
10 本地互网络都包含为在内部通讯信道和扩展互网络间传输信息
而设置的非本地传输单元。扩展互网络包含为在本地互网络间
传输信息而设置的扩展传输单元，这样信息就能可选地在本地通
讯接口的任何内部通讯信道间传输了。

在另一种实施方案中，本发明将包含多达八个带有相关本地互
15 联网络的本地通讯接口。根据本发明的动态带宽扩展这一特性，扩
展互网络在本地通讯接口的数量发生变化时将保持不变，并提供
在内部通讯信道间传输信息的能力。这样的一种实施方案可轻松地
使得带宽可扩展并与本领域的现有技术无关。在进一步的实施方案
中，通讯节点能在节点处理传输信息时调整以改变本地通讯接口
20 的数量。这样，通讯节点与本发明的一种实施方案中的互网络相结
合就可以更轻松地满足服务供应商不同的带宽需求。

如前所述，本发明可以提供增强的 QoS 特性。为提供这一特性，
根据本发明的一种实施方案的互网络将可以监视通讯信道的可用
性。更确切地说，本地互网络能包含多个为在本地传输单元和内
部通讯信道间传输信息而设置的收发器。每一个收发器都与一个相
25 关的内部通讯信道相耦合，并有一个与之相关的可用性状态以指示
那条用来传输信息的通讯信道是否可用。本地互网络还包含多个
与收发器相关的内存存储队列，还包含存储相关收发器所传输信息
的内存。本地互网络还可能会进一步包含控制单元以便能够为某
一特定的内部通讯信道设置状态而表示它不可用来传输信息，响应
30 达到可以选择内容级别的内存队列。这样，通讯节点就能降低由于
某一信道过载而导致的信息丢失或传输阻塞的可能性。此外，根据

进一步的实施方案，耦合到通讯节点中的信息被赋予一定的优先级，互联网络为表示某一信道是否具有接收一定优先级的信息的可用性而可选地包含一些控制单元以设置与之相关的状态，比如说设为高，中或低，以作为对相关内存队列达到某一可选内容级别的响应。

5 根据一种相关的实施方案，互联网络向内部通讯信道提供了一种反压（back pressure）信号，这个反压信号包含了每一个内部通讯信道是否可用的状态。本特性的进一步增强是利用那些与某一相关内部通讯信道在初始化时预留的目的地址或句柄所占用的通讯位，用它们来从本地互联网络向与之相关的本地通讯接口传输反压/可用状态。

10 根据另一种实施方案，本发明提供了增强的纠错能力。结果是本地互联网络包含了一些单元以生成从本地互联网络到本地通讯接口的冗余版的信息。如果错误检测单元检测到了传输信息中有异常，错误纠正单元就会通过冗余版的信息恢复出一个无错版的信息。根据进一步的实施方案，通讯节点包含用来通过线卡或是互联单元可使那些产生可检测异常的单元无效的控制单元。根据一种相关的实施方案，通讯节点在不降低通过此节点的信息的传输速度的条件下，通过使发生故障的线卡或是互联单元无效的方式提高容错能力。此外，为减少停机时间，一般来说通讯节点，更确切地说是互联部分

15 将会包含电路保护单元以保证通讯节点继续传输信息的条件下支持热插拔故障部件。

20 在一种相关的实施方案中，互联网络在内部是以信息信元为单位来传输信息的，其中在每一个信息信元中都包含了信息字组，而每一组信息字都是通过另一个内部通讯信道传输的。本地互联网络通过在一对信息字组传输到本地通讯接口以前，对它们进行位对位的“异或”操作来生成冗余版信息。本地互联网络也向本地通讯接口传输信息字组对的“异或”版本。作为对检测到信息字组对中任一种出现异常的响应，本地通讯接口通过对此信息字组对的“异或”版本与其非异常成员进行“异或”操作来重构一个无错的版本。

25 提高本发明的不阻塞特性的一种方法是，信息在互联网络中传输的速度将因避免信息在互联网络中传输之后对信息字组进行重新排序以形成一个完整的信元而提高。根据一种实施方案，本发明使

用“丛”(“clumping”)来避免重新排序,这样就提高了传输速度。具体的讲,互联网络会包含元素以生成“丛”或者结合多个信息信元,然后再基本上同时传输这些丛单元。进一步的增强方案是,互联网络将包含一些单元用以在那些不完整的丛单元传输之前向其中追加一些“哑”信息信元。

根据一种相关的实施方案,根据本发明的互联网络通过使用与收发器相关的存储队列来实现这一丛特性。存储队列在中间阶段存储将被传输的信息字组。互联网络将进一步包含一些检测单元,用以在多个将被包含进丛的信息信元的字组存储进队列时对它们进行检测,和一些传输单元,用以通过耦合这些字组到收发器以在稍后基本上同步的传输这些丛信息信元。

由于通讯节点通过包含辅助本地通讯接口和相关本地互联网络得到扩展,扩展后的互联网络如何选择一条高效的信息传输路径以避免不必要的延迟就显得越来越重要了。此外,在一种进一步的实施方案中,本发明提供了一种基本上的不阻塞特性。根据不阻塞特性,扩展后的互联网络可包含用于保存表示扩展后的互联网络中不阻塞本地路径数据的转发数组。扩展后的互联网络能至少使用信息信元字组中的部分目的地址做为指向转发数组的指针以选择一条不阻塞路径。在一种进一步的实施方案中,扩展后的互联网络使用多个转发数组,每一个转发数组都存储了表示扩展后的互联网络中不阻塞路径中的段的数据。接下来,扩展后的互联网络就可以利用目的地址的这些连续部分作为指向每一个前向数组的指针以选择一条不阻塞路径的每个段。

根据一种进一步的实施方案,本地互联网络的传输单元以及扩展后的互联网络的传输单元基本上是一样的,这样的话,就是可互相交换的。在这样的一种实施方案中,传输单元将具有模式选择的特性用以选择这个单元应该是在本地模式还是在扩展模式下使用。这一特性将比现有技术系统节省相当的费用开销。

另一方面,本发明包括与上述装置相应的方法。

本发明提供:

(一)一种用于在通讯节点内工作的互联网络,所述网络包括:
多个本地线卡模块,它们被配置成处理以多种速度所接收且依据

多种协议格式化的信息，

数目可选的本地互联模块，它们被连接到所述本地线卡模块且在位置上彼此靠近，所述本地互联模块中的每一个都包括本地传输单元，用于在多个本地 I/O 信道之间传输信息和在所述多个本地 I/O 信道和多个非本地 I/O 信道之间传输信息，以及

扩展互联模块，在位置上靠近所述本地互联模块，并且包括耦合装置，用于与所述非本地 I/O 信道进行电耦合，和扩展传输单元，用于在所述本地互联模块之间传输信息。

(二) 一种可动态扩展的通讯互连网络，包含：

数目可选的本地互联模块，每一个本地互联模块都包括相关的传输单元，用来将信息传输通过所述相关的本地互联模块，以及

单一的扩展互联模块，包含耦合单元，用于耦合所述数目选定的本地互联模块，和扩展传输单元，用于在所述本地互联模块之间传输信息，其中，所述数目可定的本地互联模块能够在所述扩展互联模块正传输着信息的同时加以变更。

(三) 一种用于在通讯节点内工作的互连网络，包含：

数目可选的本地互联模块，每一个本地互联模块都包括本地传输单元，用于在多个本地 I/O 信道之间传输信息和在所述多个本地 I/O 信道和多个非本地 I/O 信道之间传输信息，其中，所述互连网络适用于以信息信元来传输信息，所述信元中的每个信元都包括信息字的组，

扩展互联模块，包括耦合装置，用于与所述非本地 I/O 信道进行电耦合，和扩展传输单元，用于在所述本地互联模块之间传输信息，以及

冗余生成装置，用于通过生成一个信息信元中所包括的所述信息字组对逐位的“异或”来生成从所述互连网络向外传送中的信息的另一版本。

(四) 一种用于在通讯节点内工作的互连网络，包含：

数目可选的本地互联模块，每一个本地互联模块都包括本地传输单元，用于在多个本地 I/O 信道之间传输信息和在所述多个本地 I/O 信道和多个非本地 I/O 信道之间传输信息，

扩展互联模块，包括耦合装置，用于与所述非本地 I/O 信道进行

电耦合, 和扩展传输单元, 用于在所述本地互联模块之间传输信息, 以及

服务质量装置, 用于根据所述信息的相关优先级区分被耦合到所述本地 I/O 信道中的信息, 和确定用于接收在一个或多个所述本地 I/O 信道上具有一特定相关优先级的信息的不可用性。

(五) 一种用于在通讯节点内工作的互联网络, 所述网络包含:

数目可选的本地互联模块, 每一个本地互联模块都包括本地传输单元, 用于在多个本地 I/O 信道之间传输信息和在所述多个本地 I/O 信道和多个非本地 I/O 信道之间传输信息, 以及

至少一个扩展互联模块, 包括耦合装置, 用于与所述非本地 I/O 信道进行电耦合, 和扩展传输单元, 用于在所述本地互联模块之间传输信息,

其中, 所述互联网络适用于以信息信元来传输信息, 并且所述本地和扩展传输单元进一步包含丛装置, 用于基本上同时地传输多个所述信息信元。

(六) 一种带宽可动态扩展的通讯节点, 包含:

数目可选的本地通讯模块, 所述本地通讯模块给所述通讯节点提供带宽的动态可扩展性, 所述本地通讯模块中的每一个通讯模块都包括用于耦合进出所述节点的信息的多个外部通讯信道, 用于在所述节点之内耦合信息的多个内部通讯信道, 以及用于在所述外部通讯信道和所述内部通讯信道之间耦合信息的装置,

本地互联模块, 与所述数目可选的本地通讯模块中的每一个本地通讯模块相关, 并包括本地传输单元, 用于在所述相关的本地通讯模块的所述多个内部通讯信道之间传输信息, 和

扩展互联模块, 包含耦合装置, 用于与每一个所述本地通讯模块相耦合, 和扩展传输单元, 用于在所述本地通讯模块之间传输信息, 由此就能够在所述数目可选的本地通讯模块中的不同本地通讯模块的外部通讯信道之间传输信息。

(七) 一种带宽可动态扩展的通讯节点, 具有模块化的结构并且包含:

数目可选的本地通讯模块, 所述本地通讯模块给所述通信节点提供带宽的动态可扩展性, 所述本地通讯模块中的每一个本地通讯模

块都包括多个通讯信道，用于将信息传送到所述节点中和将信息从所述节点中传送出去，

5 本地互联模块，与所述本地通讯模块中的每一个本地通讯模块相耦合，并包括本地传输单元，用于在所述多个通讯信道之间传输信息，以及

扩展互联模块，包含耦合单元，用于与多个所述本地通讯模块相耦合，和传输单元，用于在所述多个本地互联模块之间传输信息，其中，信息可以在所述本地通讯模块的不同通讯模块的通讯信道之间传输。

10 (八)一种在适用于在通讯节点内工作的互联网络内的 I/O 信道间传输信息信号的方法，包含以下步骤：

将信息耦合到一个或多个位置上靠近的本地互联模块的 I/O 信道上，

15 响应本地目的地址，在所述本地互联模块中的一个特定本地互联模块的 I/O 信道间传输信息，

响应非本地目的地址，将信息从所述本地互联模块中的一个本地互联模块传输到位置上靠近的扩展互联模块的 I/O 信道，

响应所述非本地目的地址，将信息从所述扩展互联模块传输到所述本地互联模块的另一个特定的本地互联模块，以及

20 通过在所述多个本地互联模块中包含数目可选的所述本地互联模块来调节所述互联网络的带宽。

附图说明

25 本发明的主要问题在说明书的结论部分被特别地指出和清楚地要求了专利保护。但是，本发明既是实践的组织结构又是实践的实现方法，其中又包含了进一步的目的是优越性，因此参考以下的描述性说明以及带有对相似单元标以相似数字的附图将能最好地理解它。

图 1 说明了多个通讯网络通过本发明所说的带有互联网络的通讯节点互相联接的情况；

30 图 2 是由图 1 中说明的那种通讯节点和本发明所说的并入其中的互联网络的框图；

图 3 是一种具有代表性的图 2 中描述的那种本地互联模块与本地

线卡模块的更为详细的互联示意图;

图 4 说明与图 2 和图 3 种描述的那种互联模块连接的典型线卡模块连接的表;

5 图 5 是一种表现具有代表性的本地互联与图 2 中描述的那种扩展互联模块之间的连接的更为详细的框图;

图 6 是一种关于具有代表性的本地互联模块与扩展互联模块之间通讯信道连接的详细列表;

图 7 是进一步说明潜在的本地互联模块与扩展互联模块之间通讯信道连接的列表;

10 图 8 是图 3 中描述的那种本地互联逻辑平面的功能框图;

图 9A 和图 9B 是对高优先级队列与低优先级队列的概念说明;

图 10A 和图 10B 是描述图 9A 和图 9B 中队列的操作的流程图;

及

图 11 是图 5 中描述的那种扩展互联逻辑平面的功能框图。

15 具体实施方式

正如以上所简述的,本发明主要是针对通讯节点中的互连网络的。通讯节点是在多个信号源之间传输通讯信号的交汇点。这样,通讯节点就需要去连接使用不同通讯协议和工作在不同信息传输速度的系统。现有技术的系统通常需要维持多个网络,每一个网络用来与一种特定类型的信号源连接。相反,体现了本发明说明性的实施方案的特性的通讯节点,能够对以不同速度以及根据不同协议进行了格式化处理的进入节点的信息进行处理。举例来说,信息能够以 OC48,OC12/STM4,OC3/STM1,DS3 和 E3 速度进出通讯节点。此外,信息也能够以基于 IP 或 ATM 的格式进出通讯节点。

25 本发明的另一个重要特性是动态带宽可扩展性。一种使用本发明示出的实施方案中所提到的互连网络的通讯节点使用了模块化的设计。模块化的设计使得服务提供商可以通过增加或减少通讯节点上物理上位置接近的模块来改变通讯信道的数量。根据一种最佳实施例,这些模块将包含多个与相关互连网络耦合的 I/O 接口。在本发明的进一步的实施方案中,通讯节点使用了具有两个层次的互连网络模块结构;本地层次和扩展层次。更具体地说,多个本地互连网络模块被优先地放置在互相邻近的地方,与它们耦合的扩展互连网络

30

模块也被优先地放置在与本地互联网络模块邻近的地方。改变“插入”扩展互联网络模块中的本地互联网络模块的数量，一个服务供应商就能改变通讯节点的带宽。而且，根据一种进一步的实施方案，一个服务提供商能在通讯节点处理传输信息时连接或断开本地互联网络模块，从而提供动态的带宽可扩展性。

图 1 说明了一种根据本发明使用互联网络的通讯节点 100 的典型配置。如图所示，通讯节点 100 能够支持诸如 SONET 环 10，WWW 服务器群 20，拨号 IP 系统 30，CATV/xDSL 中 IP 系统 40，FR 中 IP 网络 50，多服务 ATM 网络 60，ATM 传输 70，和互联网 WDM 系统 80。多服务 ATM 接入 60 能包括 CBR 线路 61，ATM 上的语音线路 62，互联网接入 63，ATM 中 FR 64，和 LANs 65。

图 2 说明了使用本发明说明性的实施方案中所提到的包含互联网的通讯节点 100 的框图。通讯节点 100 使用了双层互联架构。因此，节点 100 包含一个或多个本地线卡模块 102-116，每一个模块都拥有第一级相关的本地互联模块 118-132。在节点 100 包含多个本地线卡模块的情况下，通讯节点 100 需要进一步包含第二级扩展互联模块 134，有时也被称作超连接结构。图解的扩展互联模块 134 包含三块扩展互联板 136-140。这些板是带有集成电路和其它组件的印刷电路板。根据一种实施方案，本地线卡模块 102-116 基本上是相同的。类似的，本地互联模块 118-132 在基本上也是相同的。因此，在以下的大部分篇幅中，将集中说明具有代表性的本地线卡模块 102 和本地互联模块 118；同样的说明对于剩下的本地线卡模块 104-116 和本地互联模块 120-132 也同样适用。

本地线卡模块 102 通过多个 I/O 接口传输信息进出通讯节点 100。这些 I/O 接口可能是，例如，可接受 OC48 数据流的 IP 或 SONET/SDH 端口。为兼容（grooming）低速接口，可选的接入模块 162-180 被用来提供 OC12/STM4，OC3/STM1，DS3，和 E3 端口。接入模块 162 多路复用输入数据流进入 OC48/STM16 上行线路到达本地线卡模块 102。本地线卡模块 102 通过多个 142 的千兆位以太网连接将信息耦合到相关的本地互联模块 118。本地互联模块 118 在 I/O 信道与本地线卡模块 102 之间传输信息。

本地线卡模块 102 的一个特性是它支持一定数量的数据链路层封装，并通过一种灵活的封装/解封装机制来实现。解封装机制对于新出现的封装机制是可修改的。根据一种实施方案，本地线卡模块 102 支持 IP over ATM over SONET/SDH；IP over PPP over SONET/SDH；IP over FR over SONET/SDH；IP over PPP over FR over SONET/SDH；IP over PPP over ATM over SONET/SDH；IP over MPLS over SONET/SDH；IP over SNAP 802.2；和 IP over Ethernet 2.0。线卡模块 102 还支持帧中继 DTE。

那些通讯节点领域的技术人员会理解为适应本发明而做的一些其它封装。以上所列只是为了做清楚的说明，事实上并不仅限于此。

示出的通讯节点 100 的进一步特性是它是动态带宽可扩展的。更确切地说，根据一种实施方案，通讯节点 100 能包含多达七个辅助本

地线卡模块 104-116, 并通过千兆位以太连接 144-160 与相关本地互联模块 120-132 相耦合。扩展的互联模块 134 通过多个 170 的千兆位以太连接与本地互联模块 118-132 传输信息。每一个本地互联模块 118-132 都与所有的三块扩展互联板 136-140 相耦合。本发明的另一个特性是同样的扩展的互联模块 134 可以如被辅助本地线卡模块使用一样被两个本地线卡模块使用。

图 3 是一种对于通讯节点 100 的说明实施方案 200 进行进一步地详细描述的图解的框图, 但只画了一个单独的本地线卡模块 102 与一个单独的本地互联模块 118 相耦合。由于只有一个单独的本地线卡模块 102, 通讯节点 200 就没有必要象在图 1 中那样包含第二级扩展互联模块 134 了。

本地线卡模块 102 包含八块本地线卡 202-216。本地线卡 202-216 是带有集成电路和其它组件的印刷电路板。每一块本地线卡 202-216 都有六个内部 (I/O) 端口 202a-202f, 和一个外部 SONET I/O 端口 202g。本地线卡 202 在外部端口 202g 与内部 I/O 端口 202a-202f 之间耦合信息。外部端口 202g 耦合信息进出节点 200, 内部 I/O 端口 202a-202f 与多达四十八条内部通讯线 217 相连并在本地线卡模块 102 与本地互联模块 118 之间耦合信息。一般来说, 每一个内部 I/O 端口 a-f 都包含一个千兆位以太收发器, 可以提供一个千兆位以太输入信道和一个千兆位以太输出信道。更好的是, 输入和输出信道都提供 10-位的信息。但是, 请注意在本说明中收发器这一用词, 它也指那些在分别包含接收器和发送器的结构。外部 I/O 端口 202g 对于 SONET 或 SDH 操作都是可以用软件进行配置的。这样, 对于 OC48 或 STM16 来说硬件接口是可以用软件进行配置的。SONET 和 SDH PAMS 可以自由地在接入模块 142 内进行混合使用。一个满载的本地线卡模块 102 可以有多个 SONET/SDH I/O 端口和四十八个相关的内部 I/O 端口。

本地互联模块 118 包含三块同样的互联板 218-222。互联板 218-222 是带有集成电路和其它组件的印刷电路板。每块板 218-222 在逻辑上都被分成基本上相同的两个面。举例来说, 互联板 218 包含逻辑面 218a 和 218b; 互联板 220 包含逻辑面 220a 和 220b; 同样互联板 222 包含逻辑面 222a 和 222b。通讯节点 200 利用特殊应用集成电路 (ASICs)

224-228 在互联板 218-222 之间传输信息。每一个 ASIC 224-228 在逻辑上都包含一个 a-半和 b-半。逻辑 a-半为互联板 218-222 中的逻辑 a-面服务，而逻辑 b-半为互联板 218-222 中的逻辑 b-面服务。举例来说，ASIC 224a 为逻辑面 218a 服务，ASIC 224b 为逻辑面 218b 服务。

5 图 2 和图 3 的示出实施方案着重说明了线卡 202-216 和本地互联板 218-222 在物理上的分组，及其逻辑划分本地互联面 218a, 218b, 220a, 220b, 222a 和 222c。但是，本领域的技术人员会理解线卡 202-216 和相关本地互联模块 118-132 的数目可以改变。更好的是，本地互联模块 118 中包含的本地互联板和扩展互联模块 134 中包含的扩展互联板的数目可以改变。所有的逻辑划分也是可以改变的。

10 需要注意的是，根据示出实施方案，每一块本地互联板 218-222 都包含十六个内部通讯端口（每八个与一个逻辑面相关），还包含有十六个扩展通讯端口（每八个与一个逻辑面相关）。由于每一块本地互联板在基本上都是相同的，板 218 将被用来做详细的说明。板
15 220 和 222 具有相似的结构和作用。更确切地说，本地互联板 218 有八个与 ASIC 224a 相关的内部通讯端口 0a-7a，和八个与 ASIC 224b 相关的内部通讯端口 0b-7b。本地互联板 218 还包含八个与 ASIC 224a 相关的扩展通讯端口 8a-15a，和八个与 ASIC 224b 相关的扩展通讯端口 8b-15b。每个内部和扩展通讯端口包括提供千兆位以太输入信道和千兆位以太输出信道的以太收发器。每一个内部和扩展通讯端口
20 0a-7a 和 0b-7b 都与线卡 202-216 的一个内部通讯端口 a-f 相耦合，并在本地线卡模块 102 与本地互联模块 118 之间传输信息。类似的，每一个线卡 202-216 的内部通讯端口 a-f 都与互联板 218-222 的一个内部通讯端口 0a-7a 和 0b-7b 相耦合。

25 图 4 展示了表 300 以说明线卡 202-216 中的内部通讯端口与互联板 218-222 中的内部通讯端口之间的相互联接。列 302 显示了线卡的标志数字 202-216。列 304 列出了（互联面#，端口#）形式的有序对组，这些有序对确定了互联模块 118 中的端口 0a-7a 和 0b-7b 中的哪一个与线卡 202-216 中的每一个内部端口 a-f 相耦合。举例来说，表
30 300 中的第一行表示线卡 202，内部端口 a 与互联面 218a 的内部端口 0a 相耦合。类似的，线卡 202 的内部端口 b, c, d, e 和 f 分别与面 218b 的内部端口 0b，面 220a 的内部端口 0a，面 220b 的内部端口 0a，面

222a 的内部端口 0a 和面 222b 的内部端口 0b 相耦合。换句话说，线卡 202 的内部通讯端口与互联面 224a, 224b, 226a, 226b, 228a, 和 228b 中的“零数”端口相耦合。与此相似，线卡 204 的内部通讯端口与“一数”端口（例如 1a）相耦合。线卡 206 的内部通讯端口与“二数”端口（例如 2a）相耦合。线卡 208 的内部通讯端口与“三数”端口（例如 3a）相耦合，以此类推。

根据一种示出的实施方案，下面如表 1 所显示的，通讯节点 200 在本地互联模块 118 与相关的本地线卡模块 102 之间以 64-字节的信元为单位传输信息。

如表 1 所示，64-字节的信元被分成 16-字节长的组。逻辑面 218a 传输 8-字，每字 2-字节。逻辑面 218b, 220a 和 220b 也是同样如此。在第一和第三 16-字节组（即由逻辑面 218a 和 220a 传输的组）中最低的两个字节（LSBs）被用作传输目的线卡的地址/编号。在第二和第四 16-字节组（即由逻辑面 218b 和 220b 传输的组）中的 LSBs 被用作传输互联寻址和流量控制信息。板 222 提供了纠错和冗余信息。更具体而言，逻辑面 222a 提供了对逻辑面 218a 与逻辑面 220a 之间传输的信息进行位对位的“异或”（ $\oplus$ ）操作。逻辑面 222b 提供了对逻辑面 218b 与逻辑面 220b 之间传输的信息进行位对位的“异或”操作。在表 1 的示出的实施方案中，一个字节包含 8-位而一个字则包含 2-字节。但是，那些本领域的技术人员会理解可以选择其它的字节与字的规范。

表 1

平面	字 7	字 6	字 5	字 4	字 3	字 2	字 1	字 0
218a	数据 7-0	数据 6-0	数据 5-0	数据 4-0	数据 3-0	数据 2-0	数据 1-0	目标地址
218b	数据 7-0a	数据 6-0a	数据 5-0a	数据 4-0a	数据 3-0a	数据 2-0a	数据 1-0a	流程控制
220a	数据 7-1	数据 6-1	数据 5-1	数据 4-1	数据 3-1	数据 2-1	数据 1-1	目标地址
220b	数据 7-1a	数据 6-1a	数据 5-1a	数据 4-1a	数据 3-1a	数据 2-1a	数据 1-1a	流程控制
222a	数据 7-0 $\oplus$ 数据 7-1	数据 6-0 $\oplus$ 数据 6-1	数据 5-0 $\oplus$ 数据 5-1	数据 4-0 $\oplus$ 数据 4-1	数据 3-0 $\oplus$ 数据 3-1	数据 2-0 $\oplus$ 数据 2-1	数据 1-0 $\oplus$ 数据 1-1	目标地址
222b	数据 7-0a $\oplus$ 数据 7-1a	数据 6-0a $\oplus$ 数据 6-1a	数据 5-0a $\oplus$ 数据 5-1a	数据 4-0a $\oplus$ 数据 4-1a	数据 3-0a $\oplus$ 数据 3-1a	数据 2-0a $\oplus$ 数据 2-1a	数据 1-0a $\oplus$ 数据 1-1a	流程控制

线卡 202-216 中的每一块都有六个内部的千兆位以太端口 a-f，它

们分布于三块互联板 218-222 之间，根据表 1 中定义的格式，六个内部的千兆位以太端口 a-f 将提供 3-Gbs 的可用带宽。更确切地说，板 222 并没有被用做有效载荷带宽，而是提供了冗余和纠错信息，这样就剩下了 4-Gbs 的带宽。传输信元中每六十四字节中有四个字节（逻辑面 218b 与 220b 中的 LSBs）被用作互联寻址和流量控制信息，这样就剩下了 3.5-Gbs 的带宽。然后，剩下的每五十六字节中有四个字节（逻辑面 218a 与 220a 中的 LSBs）被线卡 202-216 用作目的句柄/地址，这样就剩下了 3-Gbs 的带宽。这保证了通讯节点 200 能够稳定的提供 OC-48（2.4Gbs）的传输速率。

下面的表 2 说明了从线卡 202-216 到本地互联面 218a, 218b, 220a, 220b, 222a 和 222b 之间传输的信息的标准信息信元格式。

如表 2 中第一列所示并如在表 1 中前面已经说明的，信息是以 8-字/16-字节的组为单位传输的。每一个逻辑面 218a, 218b, 220a 和 220b 接收 16-字节组。逻辑面 222a 接收面 218a 和 220a 的“异或”结果，而逻辑面 220b 则接收面 218b 和 220b 的“异或”结果。逻辑面 218a, 218b, 220a 和 220b 中的字节 2-16 提供所传输的数据。

表 2

字节	面 218a	面 218b	面 220a	面 220b	面 222a	面 222b
0	目标低	0 <u>X</u>	目标低	0 <u>X</u>	目标低	0 <u>X</u>
1	目标高	P <u>有效载荷</u>	目标高	P <u>有效载荷</u>	目标高	P <u>有效载荷</u>
2	数据 0	数据 1	数据 2	数据 3	218a+220a	218b+220b
3	数据 4	数据 5	数据 6	数据 7	218a+220a	218b+220b
15	数据 52	数据 53	数据 54	数据 55	218a+220a	218b+220b

上表中的“X”表示从线卡到本地互联模块 118 的一个 XOFF。信息信元通过分离的面 218a-222b 经过本地互联模块 118，其中略有区别。举例来说，a-面包含了线卡的目的地址。b-面允许 14-位的额外“有效载荷”未改变而通过。上面提到的有效载荷字节典型地用最高位（“P”位）作为字 0 和字 1 的奇偶校验位。即使是在面 222a 上，“P”位也是覆盖了在它前面的 15-位，而不是对面 218a 和 222a 的奇偶校验。

线卡的目的地址是指定信息信元将被发送到的目的线卡（对单点发送信息）或线卡集（对多点发送信息）的地址或句柄。

下面表 3 说明了对于在本地传输的单点发送信息的优选的目的地址格式，而表 4 则说明了对于在本地传输的多点发送信息的优选的目的地址格式。

表 3

字节	7	6	5	4	3	2	1	0
0	APS	0	线卡模块			线卡指定		
1	P	0	0	0	0	0	1	Pri
2-15	信元数据							

表 4

字节	7	6	5	4	3	2	1	0
0	多点传送 ID 的最低 8 位							
1	P	1	多点传送 ID 的最高 6 位					
2-15	信元数据							

表 3 中的字节-0 提供了一个“APS”位，它是线卡模块指定域和线可指定域。如图 3 所示，根据示出的实施方案，有八个潜在的线卡 202-216。每一个都被指定一个 3-位编码 000-111。表 3 中的字节-0 中的位 3，4 和 5 提供了这一编码。如图 2 所示，根据示出的实施方案，通讯节点 100 有八个潜在的线卡 102-116。每一个每个线卡模块 102-116 都被指定一个 3-位编码 000-111。表 3 中的字节-0 中的位 0，1 和 2 提供了这一编码。字节 2-15 提供了所传输的数据。

“P”位提供了对字节 1 的奇偶校验。

根据优选实施方案，传输到单个目的线卡的单点发送信息信元在需要较高优先级流量时把表 3 所示的“Pri”位置为一。如果表 1 中的“APS”位被置位的话，信元将被同时发送到指定线卡(n)和(n+1)线卡。

参考表 4，多点发送 ID 是图 8 中的 16k×9-位 RAM748 的一个地址。此 14-位地址取回一个 9-位的值。低 8-位是每一位表示一块线卡 202-216 的掩码。最高位表示优先级；置为逻辑一表示高优先级或置为逻辑零表示低优先级。如果节点包含不止一块本地线卡底板 102，

那么这个 8-位掩码中的每一位都代表了一块特定的本地线卡底板。

如上面的表 1 所示，在操作中，通讯节点 200 通过不同的内部互
联信道传输每一个 16-字节组。在此用图 3 举例来说，假设线卡 202-
216 中的每一块都有一个相应的地址，而信息通过外部端口 202g 进
5 入线卡 202。进一步假设进入的信息包含线卡 208 的目的地址。如图
2 所示，内部端口 202a 把第一个 16-字节组耦合到互联板 218 的内部
端口 0a。内部端口 202b 把第二个 16-字节组耦合到互联板 218 的内
部端口 0b。内部端口 202c 把第三个 16-字节组耦合到互联板 220 的
内部端口 0a。内部端口 202d 把第四个 16-字节组耦合到互联板 220
10 的内部端口 0b。板 218 的内部端口 0a 把第一个 16-字节组耦合到 ASIC
224a。板 218 的内部端口 0b 把第二个 16-字节组耦合到 ASIC 224b。
板 220 的内部端口 0a 把第三个 16-字节组耦合到 ASIC 226a。板 220
的内部端口 0b 把第四个 16-字节组耦合到 ASIC 226b。

ASIC 224a 处理线卡的目的地址并把第一个 16-字节组导向到板
15 218 的内部端口 3a。ASIC 224b 处理线卡的目的地址并把第二个 16-
字节组导向到板 218 的内部端口 3b。ASIC 226a 处理线卡的目的地址
并把第三个 16-字节组导向到板 220 的内部端口 3a。ASIC 226b 处理
线卡的目的地址并把第四个 16-字节组导向到板 220 的内部端口 3b。
板 222 在第一组和第三组之间以及第二组与第四组之间生成位对位的
20 “异或”值。ASIC 228a 处理线卡的目的地址并把第一组和第三组合
成的“异或”值导向到板 222 的内部通讯端口 3a，ASIC 228b 处理线
卡的目的地址并把第二组和第四组合成的“异或”值导向到板 222 的
内部通讯端口 3b。板 222 依次把第一，第二，第三和第四组分别耦
合到目的线卡 208，内部端口 208a-208d。类似的，板 222 把“异或”
25 版的字节组耦合到内部端口 208e 和 208f。目的线卡 208 然后进行各
种类型的错误检查，诸如对多个信元进行简单奇偶校验，8B10B 不
一致校验和 CRC 校验。如果线卡 208 检测出了一个字符错误，软件
会警告线卡去用“异或”版本来重新得到所传输信息的无错版本。

根据一种进一步的实施方案，线卡模块 102 能够确定一个检测到的
30 的错误是由一个发生故障的线卡 202-216 引起的或是由一个发生故障
的互联板 218-222 引起的。在错误是由一个发生故障的线卡引起的
时，互联模块将停止此板的操作。在错误是由一个发生故障的互联

板 218-222 引起的时, 检测到这个错误的线卡会向有可能发生错误的互联板 218-222 发送信号以使自己离线。为了不影响带宽, 互联板 222 会自动替换互联板 218 或 220 的位置, 直至发生故障的板被更换。根据一个进一步的特性, 发生故障的板是可以热插拔的。

5 如上面所讨论的, 根据一种优选的实施方案本发明的一个特性是, 通讯节点 100 可以包含辅助的线卡模块 104-116 以具有动态的带宽可扩展性它有辅助的相关本地互联底板 120-160。根据一种优选的实施方案, 线卡模块 102-116 的模块化结构和本地互联模块 118-132 的模块化结构, 以及扩展的互联模块 134 提供了可扩展的特性。更具体地说, 如图 2 所示, 根据示出的实施方案 100, 第一块机械底板 101 10 包含了 8 个本地线卡模块 102-116。类似的, 第二块机械底板 103 包含了 8 个本地互联模块 118-132。进一步的, 第三块机械底板 105 包含了扩展互联网络 134。第一块机械底板 101 通过通讯信道 142-160 与第二个机械底板 103 电气耦合。第二个机械底板 103 在电路上通过 15 通讯信道 170 与第二个机械底板 103 相耦合。通讯节点 100 在模块接口上使用设计上支持“热插拔”的连接器的连接器。这些连接器使本地线卡模块 102-116, 本地互联模块 118-132 和扩展互联板 136-140 可以和它们分别的机械底板 101, 103 和 105 在通讯节点 100 加电和传输信息时连接和断开(即“热插拔”)。这样, 当一个服务提供商需要 20 而外带宽时, 辅助本地线卡模块 102-116 以及相关辅助本地互联模块 118-132 就可以分别“插入”底板 101 和 103。

图 5 是一种更详细的表示一个典型的本地互联模块 118 与一个扩展互联模块 134 之间连接的原理框图 400。图 4 的实施方案说明了通讯节点 100 的动态的带宽可扩展性。如上所述, 本地互联模块 118 和 25 可选的本地互联模块 120-132 在基本上是相同的。从而, 下面的说明性描述将集中在本地互联模块 118 和扩展互联模块 134 上。

如图 5 所示, 并且如图 3 中所详细讨论的, 本地互联模块 118 包含 30 三块互联板 218-222。互联板 218-222 中的每一块在逻辑上都被分成 a-面和 b-面。a-面和 b-面在基本上是相同的且共享一个在逻辑上也被分成 a-面和 b-面的 ASIC。每一个 ASIC 224-222 在提供其它功能的同时, 也提供在与特定互联板 218-222 相关的内部通讯信道之间的路由和交换。根据示出的实施方案, 每一个 ASIC 224-222 也可能会

提供存储队列，一个控制处理器，控制寄存器和状态寄存器。它也控制包含在每一块互联板 218-222 上的千兆位以太 I/O 接口。ASIC 224-222 还进一步为存储指示交换/路由线路中基本上不阻塞路径信息而提供的一个或多个指针队列。

5 每一块互联板 218-222 都包含十六个内部 I/O 端口 0a-7a 和 0b-7b。内部 I/O 端口 0a-7a 和 0b-7b 提供千兆位以太接口。如图 3 所示，这些千兆位接口与一个相关本地线卡模块 102 的内部通讯信道相耦合。每一块互联板 218-222 也提供了十六个扩展 I/O 端口 8a-15a 和 8b-15b。每一个扩展 I/O 端口 8a-15a 和 8b-15b 提供了一个千兆位输入信
10 道和一个千兆位输出信道。扩展 I/O 端口 8a-15a 和 8b-15b 耦合本地互联模块 118 与扩展互联模块 134 之间的信息。

扩展互联模块 134 包含三块基本上相同的扩展互联板 136-140。每一块板 136-140 除其它组件外都包含一百二十八个千兆位以太收发器。每一块板 136-140 还分别包含四个 ASICs 402-408, 410-416 和
15 418-424。ASICs 402-424 与 ASICs 224-228 在基本上是相同的。但是，ASICs 402-424 在模式上被选择工作在扩展互联模式，而不是 ASICs 224-228 的本地互联模式。与 ASICs 224-228 一样，每一个 ASIC 402-424 在逻辑上都被分成 a-半和 b-半。每一半都包含十六个千兆位以太 I/O 端口，每一个端口包含一个千兆位输入信道和一个千兆位输出信
20 道。十六个千兆位以太端口的每一个都与扩展互联板中的一个千兆位收发器相耦合。

举一个特殊的例子，图 5 中的板 136 包含 ASICs 402-408。ASICs 402 被分成逻辑半 402a 和 402b。类似的，ASICs 404 被分成逻辑半 404a 和 404b；ASICs 406 被分成逻辑半 406a 和 406b；ASICs 408 被分成
25 逻辑半 408a 和 408b。ASIC 402 在 402a 这一半包含千兆位以太端口 0a-15a，在 402b 这一半包含千兆位以太端口 0b-15b；端口 0a-15a 与板 136 的收发器 0-15 相耦合，而端口 0b-15b 与收发器 16-31 相耦合。ASICs 404-408 千兆位端口 0a-15a 和 0b-15b 则连续地与剩下的收发器 32-127 相耦合。ASICs 410-416 和 418-424 的千兆位端口与板 138 和
30 140 的一百二十八个收发器分别耦合，其方式与板 136 和 ASICs 402-404 所描述的相同。

图 6 是表 500 说明了本地互联模块 118 和扩展互联模块 134 之间

的通讯路径。列 502 指定了本地互联端口的名称。列 504 指定了本地
 互联板 218 和扩展互联模块 134 之间的通讯路径。类似的，列 506 指
 定了本地互联板 220 和扩展互联模块 134 之间的通讯路径；列 508 指
 定了本地互联板 222 和扩展互联模块 134 之间的通讯路径。从图 4 和
 5 可以看出，板 218 的通讯信道与板 136 相耦合。板 220 的通讯信道
 与板 138 相耦合，而板 222 的通讯信道与板 140 相耦合。从图 5 还可
 以看出，本地互联模块 118 的 I/O 接口与 ASICs 402-404 的端口零和
 八相耦合。

图 7 是表 600 说明了图 1 中的每一个本地互联模块 118-132 的扩
 展互联端口的分配方法。列 602 指定了特定的本地互联模块 118-132。
 列 604 指定了扩展互联模块 134 中的端口分配。如图所示，本地互联
 模块 120 与端口一和九相耦合；模块 122 与端口二和十相耦合；模块
 124 与端口三和十一相耦合；模块 126 与端口四和十二相耦合；模块
 128 与端口五和十三相耦合；模块 130 与端口六和十四相耦合；模块
 132 与端口七和十五相耦合。

下面的表 5 指定了从本地互联模块 118-132 传输到扩展互联模块
 134 的单点发送信息的目的地址的优选格式。类似的，表 6 指定了从
 本地互联模块 118-132 传输到扩展互联模块 134 的多点发送信息的目
 的地址的优选格式。

表 5

目标高								目标低							
7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
P	0	有效	丛	0	0	1	Pri	APS	0	模块			主		

参考表 5，最高字节（MSB）的位 7 是奇偶校验位，它是代表目
 的地址的前 15 位的奇偶校验值。位 5 是“有效”位。当目的地址有
 效时“有效”位被置位。位 4 是“丛”位。当信元的组合或丛有效时
 “丛”位被置位。丛是本发明为减少对要顺序传输的信息做重排序
 的一个特性。参考图 7，在下面我们将进一步详细讨论。位 7 是“Pri”
 位。当有高优先级的信息流时“Pri”位被置为一。LSB 的位 7 是“APS”
 位。“APS”位只是被简单的从扩展互联模块 134 传输到本地互联模
 块 118-132。LSB 的位 3-5 为目的线卡模块提供了 3-位指定编码 000-
 111。LSB 的位 0-3 为特定线卡提供了 3-位指定编码 000-111。

表 6

目标高								目标低							
7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
P	1	多点传送 ID (非 0X3FFF)													

5 优选地，表 6 中的“多点传送 ID”被传送到扩展互联模块 134 以被译码。如果信元的有效载荷数据部分为 0x3FFF，信息信元将被认为是无效的。

10 表 7 和 8 分别描述了一个从扩展互联模块 134 传输到本地互联模块 118-132 的信息信元，用于单点和多点发送目的地址的优选格式。如上所述，“P”位提供了对目的地址的奇偶校验。当目的地址有效时“有效”位被置位。当有高优先级的信息流时“Pri”位被置位。当信元被同时发送到线卡 (n) 和 (n+1) 线卡时“APS”位被置位。位 0-3 提供了信元将被发送到的线卡的指定编码。

表 7

目标高								目标低							
7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
P	0	有效					Pri	APS							主

15 表 8

目标高								目标低							
7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
P	1	多点传送 ID (非 0X3FFF)													

20 图 8 是图 5 中所示的一种本地互联板的功能框图。板 218 包含一个本地互联模块 118 使用的 ASIC 224。优选地，ASIC 224 与扩展互联模块 134 使用的 ASICs 402-424 是相同的。由于所有的本地互联板都最好是相同的，所以下面的讨论将假定板 218 是一个本地互联模块 118 和图 2 中本地线卡模块 102 的接口的本地互联板的一个例子。但是，本领域的技术人员可以理解 ASIC 224 与 ASICs 402-424 不一定相同。

25 如图 8 所示，互联板 218 包还千兆位收发器组 704 和 708，存储单元 710，控制器 712 和控制寄存器 714。千兆位收发器 708 提供了与图 3 中本地线卡模块 102 的内部通讯信道相耦合的千兆位 I/O 端口 0a-7a 和 0b-7b。千兆位收发器组 704 提供了与图 5 中扩展互联模块 134

的扩展通讯信道相耦合的千兆位 I/O 端口 8a-15a 和 8b-15b。

收发器组 704 和 708 中的每一个都通过相关的输入和输出移位和保持寄存器与 ASIC 224 相耦合。更具体地说，收发器组 704 通过连线 716 与输入移位和保持寄存器 714 相耦合，通过连线 720 与输出移位和保持寄存器 718 相耦合。收发器组 704 通过连线 724 与输入移位和保持寄存器 722 相耦合，通过连线 728 与输出移位和保持寄存器 726 相耦合。

ASIC 224 也包含了双端口 RAM 730 用来存储与流量控制信息相关的不同的栈和队列 731。流量状态 733 根据接收信息的特定线卡的可用性存储了可用的状态。RAM 730 在中间存储板 218 传输过来的信息。移位和保持寄存器 714 和 716 分别通过连线 732 和 734 与双端口 RAM 730 相耦合。移位和保持寄存器 722 和 726 通过连线 736 和 738 与双端口 RAM 730 相耦合。双端口 RAM 730 也通过连线 742 与目的栈 740 相耦合。具有九十六个目的队列 740 在中间存储特定数据在 RAM 730 中的存储地址。队列 740 为方便寻址优选地使用了多个栈。但是，其它存储结构也是可以使用的。

如前面发明概述所述，以及下面将要进一步详细讨论的，根据一种优选的实施方案，本发明使用了多个内存存储队列/缓冲区用来提高传输信息的效率。应该注意的是术语队列和缓冲区是可以互换的。双端口 RAM 730 为每一个收发器组 704 和 706 提供了一个输出队列。更具体地说，要发送到本地互联模块 102 上线卡 202-204 的与板 218 耦合的信息信元，首先会被写入其地址已被写入一个输出队列的缓冲区内。自由列表内存 742 提供了一系列可用缓冲区内地址。双端口 RAM 730 中 1536 个缓冲区中的每一个都有一个引用计数器 744。引用计数器包含输出队列的个数，各个缓冲器内容被发送到输出队列。引用计数器 744 降低了从关联缓冲区读取信息的响应时间。当引用计数器变为零时，缓冲区的地址被返回到自由列表 743。这样，ASIC 224 就能跟踪与每一个收发器相关联的可用缓冲区的位置。被写入缓冲区内信息接着被传输到了输出移位和保持寄存器 720 或 728 中的之一，并被保存在那里直到可进行目的地址查找的时间片到来，可对自由列表内存 742 进行读操作，也可对缓冲区内进行写操作，以及对输出队列进行写操作。

根据一种优选实施方案，本发明提供了增强的 QoS 特性。为此，队列 731 包含 QoS 队列。QoS 队列，例如象在概念图 9A 和 9B 中所示的，可以有多重水位标记级别；这些级别与不同的优先级相关。举例来说，图 9A 中的高优先级队列 900 有 906 和 908 两个水位标记。在范围 902 内，队列 900 报告它的状态是“停止-空”，表示与队列 900 相关的 I/O 信道已经准备好接收任何优先级的信息。如图 10A 中的流程图所描述的，在操作过程中在范围 904 内，队列 900 报告它的状态是“停止-低” 1002，表示与队列 900 相关的 I/O 信道已经准备好接收“中”或更高优先级的信息，如 1004 所示。当队列 900 被填充到 906 级时，如 100B 所示它将报告它的状态是“停止-所有”，如 1006 所示。如 1008 所示，这时表示与它相关的 I/O 信道都不可用。如 1006 所指示的，如果“停止-低”队列 900 的水位标记 908 没有达到，它将不能接收任何优先级的信息。

如图 9B 中所描述的象队列 910 这样的低优先级队列，其操作已经被图 10A 中的流程图说明，可以包含三个水位标记 918, 920 和 922。如 1012 所示，队列 910 在范围 912 内报告它的状态是“停止-空”。如 1014 和 101B 所示，它在范围 914 内报告它的状态是“停止-低”。如 1018 和 1020 所示，它在范围 916 内报告它的状态是“stop-medium”，而如 1022 和 1024 所示，队列 910 在达到水位标记 918 时报告它的状态是“停止-所有”。

象队列 900 这样的高优先级队列可以使相关的线卡通过低和中优先级信息流而在一块线卡的低优先级的信息流不抑制另一块线卡的中优先级信息流。

象队列 910 这样的低优先级队列可以使相关的线卡可以通过低，中和高优先级信息流而在一块线卡的低和中优先级的信息流不抑制另一块线卡的高优先级信息流。它也会防止一块线卡的低优先级的信息流抑制另一块线卡的中和高优先级信息流。

为了有效的管理具有不同优先级的信息，双端口 RAM 730 优选地为六十四个低优先级的单点发送队列提供了存储器；通讯节点 100 中的每一块可能的线卡都有一个。RAM 730 也为六十四个高优先级的单点发送队列提供了存储器；本地互联模块和辅助本地互联模块中的每一块线卡都有一个，并有一个额外队列。多点发送信息流优

选的使用四个低优先级和四个高优先级队列。

此外，扩展互联模块 134 中的每一个面都使用八个高优先级的单点发送队列；潜在本地互联模块 118-132 中的每一个都有一个。每一个扩展互联逻辑面也都使用八个高优先级和八个低优先级的多点发送队列；潜在本地互联目的模块 118-132 中的每一个也都有一个。

一个相关的组件，队列深度逻辑电路 746 维护了本地模块 102 中所有线卡 202-216 的状态。这些状态提供了每一块线卡 202-216 是否可以接收不同优先级信息的信息。

本发明的示出实施方案的另一个特性，是节点 100 在扩展互联模块 134 与本地互联模块 118-132 中的每一块线卡之间传输控制状态(有时也被称为反压状态)的方式。根据一种优选实施方案，本发明利用了先前为目的地址预留的信息信元中的位。这些位曾在表 1 中被称为 b-信道的“流量控制”字。

流量控制信息通过 b-信道中的最低位字在本地互联模块 118-132 与扩展互联模块 134 中的每一块线卡之间被传输。这些位被包含在主信道的目的地址字中的奇偶校验位的奇偶校验计算之中。这种格式已在前面的表 1 中基本说明了，相对于本地互联面 218a 来说是字 0，相对于本地互联面 218b 来说也是字 0。这些流量信息最好不要在所有连接中都重复。如下面的图 9 所示，也参考本地互联面 218a 和 218b 和扩展互联板 136，流量控制信息以双信元方式发送。

更确切地说，表 9 中的列 1 以 (ASIC 引用名称，ASIC 端口名称) 的格式列出了扩展互联端口。列 2 列出了本地扩展互联 218b 的引用端口名称。类型 0 和类型 1 标识了字节中所含的信息 (例如，如果本地互联端口 8b 接收了一个类型 0 的字节，则此字节将包含低 6，低 5，低 4，低 3，低 2，低 1 和低 0 流量控制信息)。当相关的低优先级队列不满也就是可以接收数据时，每一个低 0- 低 63 就被置位。类似的，当相关的高优先级，多点发送高优先级和多点发送低优先级队列有可以接收信息的可用空间时，高 0-高 7，MC 高和 MC 低位被置位。

30

表 9

外部 互联 端口	本地 互联 端口	D7	D6	D5	D4	D3	D2	D1	D0
402b, 0	8b	类型 0	低 6	低 5	低 4	低 3	低 2	低 1	低 0
404b, 0	9b	类型 0	低 13	低 12	低 11	低 10	低 9	低 8	低 7
406b, 0	10b	类型 0	低 20	低 19	低 18	低 17	低 16	低 15	低 14
408b, 0	11 b	类型 0	低 27	低 26	低 25	低 24	低 23	低 22	低 21
402b, 8	12b	类型 0	低 34	低 33	低 32	低 31	低 30	低 29	低 28
404b, 8	13b	类型 0	低 41	低 40	低 39	低 38	低 37	低 36	低 35
406b, 8	14b	类型 0	高 0	低 47	低 46	低 45	低 44	低 43	低 42
408b, 8	15b	类型 0	高 7	高 6	高 5	高 4	高 3	高 2	高 1
402b, 0	8b	类型 1	低 52	低 51	低 50	低 49	低 48	MC 高	MC 低
404b, 0	9b	类型 1	低 57	低 56	低 55	低 54	低 53	MC 高	MC 低
406b, 0	10b	类型 1	低 62	低 61	低 60	低 59	低 58	MC 高	MC 低
408b, 0	11b	类型 1	低 3	低 2	低 1	低 0	低 63	MC 高	MC 低
402b, 8	12b	类型 1	低 10	低 9	低 8	低 7	低 6	低 5	低 4
404b, 8	13b	类型 1	低 17	低 16	低 15	低 14	低 13	低 12	低 11
406b, 8	14b	类型 1	高 0	低 23	低 22	低 21	低 20	低 19	低 18
408b, 8	15	类型 1	高 7	高 6	高 5	高 4	高 3	高 2	高 1

5 尽管高优先级和一些低优先级的流量控制信息在两个循环中被重复，还是会有一些相关的延迟。这样，为避免队列溢出，水位标记级别被编程为一个考虑了潜在延迟的级别。举例来说，如果流量控制延迟使四个信元停止输入信息， $(4 \times 16) - 4$ 为置将在水位标记之上被保留以避免溢出。它的起因是每十六个本地输入端口潜在地对准了队列达四个信元时间。这样，它将在那个时间间隔内把四个信息信元耗尽。与此相对照，水位标记以下的空间只需要 $(1 \times$ 流量控制延迟) 来避免溢出。结果，一种优选实施方案把水位标记阈值级别设置为一百九十二个字节中的十二到十八字节之间。

15 根据上面讨论的结构和协议，互连网络支持恒定位速率 (CBR)，可变实时位速率 (VBR-rt)，可变非实时位速率 (VBR-nrt) 和非特定的位速率 (UBR) 的 QoS 种类。互连网络可以作为一个基于种类的 ATM 交换器来工作。这样，信息流就会根据虚电路中的服务种类

组成队列以被传输。但是，整形和策略是基于每个虚拟电路（per-virtual-circuit）工作的。互联网络也支持诸如差异服务模型的 IP 网络 QoS 特性。

5 上面还提到，本发明的一种优选实施方案使用“丛”来提高互联网络中传输信息的速率。典型的，在使用现有技术的系统中，部分的通讯可以以不同的速度通过互联网络，这样就会以错误的顺序到达同一个目的地。对靠后面传输的信息进行重排序会浪费宝贵的时间，而且会有接收次序颠倒的信元的可能性。

10 所以，根据本发明的一种优选实施方案，扩展互联网络 134 包含“丛”单元或组合了多个信息信元，并且在基本上同时发送这些丛信元。

更确切地说，当一个四个单点发送信元组在一个单独的队列中可用时会被队列深度逻辑 746 检测到。作为对检测到四个单点发送信元组在一个单独的队列中的响应，队列深度逻辑 746 会通知双端口 RAM 15 731 与目的栈 740 协同工作，以传输那检测到的四个信元丛到移位和保持寄存器 726，接下来就可以由以太收发器 708 来同时传输。丛信元是在一起传输的，它们在很短的时间内到达目的地从而避免了重排序。

20 根据一种进一步的实施方案，可编程的等待计时器当包含在丛中的第一个信息信元到达时开始计时。如果计时器在整个丛完整到达之前到时，它会触发已经准备好发送的信元与 4-N 无效信元结合，这里的 N 是这个丛中所缺少的信元数目。

25 多点发送信元跨越路径组合在一起。当多点发送信息流可以由至少四个不同的路径发送时，它被认为是可以发送的。而在单点发送信息流的情况下，在任一个给定多点发送队列上的可编程的等待计时器会仿真的提供合格的多点发送信息流。关于多点发送队列的可编程的水位标记阈值也会人工地提供合格的多点发送信息流。只要多点发送信息流是符合发送条件的，“QInfo”信元就从剩下的连接或从扩展互联模块 134 上被发送。

30 ASIC 224 也包含了翻译内存 748。翻译内存 748 为扩展互联模块 118 上的路径分段提供了存储器，如果这样的模块包含在本系统内。如图 2 中所示的节点情况，翻译内存 748 优化地包含了九个逻辑存储

区域；每个本地互联模块 102-116 一个，扩展互联模块 118 有一个。扩展互联存储区域被设置成关于目的线卡和优先级的位图。目的地址电路 750 接入翻译内存 748，多点发送位图寄存器 752 则接收接入的信息。

5 本发明的一个特性是对一个特定的本地互联模块上的本地互联板之间进行同步，以及对扩展互联板 138-140 之间进行同步。因为每一块板都是独立的，所以尽管它们从同一个源获得时钟频率，还是需要一些信号来建立和保持板之间的同步。时间片同步在基本上要求在同一个模块中的板之间和面之间有同样的“时间零”基准。信元同步使一个本地互联模块可以设置它的时间片零基准，这样它发送的信息信元接可以在安全的时间到达扩展互联模块 134。CPU 同步使得某一 CPU 的写操作对一个模块的所有面在同一时间发生。

10 面同步逻辑 752 为在扩展互联模块 118 上的每一个面之间同步一定的读和写操作而提供了所需的信号 754。如果是在本地模式下工作，或对于模块 134 的辅助扩展互联面是在扩展模式下工作，控制寄存器 758 通过连到扩展互联模块 118 上其它面的连线 760 来提供流量控制信息。

15 面到面的信元同步是通过 cellok 面间连接 761 实现的。被声明的 cellok 信号 761 说明相关的面接收到了包含 2-字节目的地址的有效或无错误信元头。根据示出的实施方案，每一个面输出 16 个 cellok 信号 761 并且输入 32 个 cellok 信号 761。每一个 cellok 输出，N，表示 a-端口和 b-端口都有有效信元头。

20 对于需要传送的信元，所有工作的面声明它们各自的 cellok 信号 761。如果一个面声明它的 cellok 信号 761 而其它的面并不这样做，错误将被记录在 CPU 可访问的寄存器 758。如果一个面发生了故障，系统将有能力命令工作的面忽略掉这个故障面。这样，一个单独的故障面不会降低受到影响的本地或扩展互联模块传输信息的速率。

25 本地互联模块 118-132 和扩展互联模块 134 时用在基本上相同的 ASICs。所以，ASIC 224 包含用于选择 ASIC 224 是按一个本地互联电路工作还是按一个扩展互联电路工作的模式选择 756。如图 3 和 5 所示，在本地模式下，端口 0a-7a 和 0b-7b 与本地线卡相连，端口 8a-15a 和 8b-15b 与扩展互联模块 134 相连。或者，在扩展模式下所有的端

口 0a-15a 和 0b-15b 与本地互联面相连，就象图 5 中所示的 218-222。

再次参照图 8，图示 ASIC 224 的另一个重要的特性是包含在计时器，计数器，控制寄存器 758 中的一个“时间片计数器”。时间片计数器反复地计数 0-15。每一个端口 0l-15a 和 0b-15b 被赋予一个时间片计数。每次当时间片计数 0-15 与一个端口号符合时，一个确定是否有信元从那个端口发送的检查将被执行。如果有的话，此信元将被从 RAM 730 拷贝到移位和保持寄存器 718 或 726 以备发送。如果没有信元将被发送，那么一个流量控制信元将被发送。根据一个示出的实施方案，一个普通的时间片计数器被 a-端口和 b-端口所使用。

如上所述，板 218 也包含控制器 712 和内存 710。内存 710 存储板 218 的控制编码。同样地，它提供状态，指针和通讯接口的启动初始化。控制器 712 提供了多个常规处理器的功能。

在本地互联板的条件下，扩展互联板在逻辑上被分成在基本上相同的 a-面和 p-面。这样，为了图解方便，图 11 使用了图 5 中扩展互联板 316 的 a-面 136a 这一用词。

面 136a 包含四个 ASICs 402a，404a，406a 和 408a。ASICs 402a-408a 在基本上与图 8 中的 ASIC 224 是相同的。每一个 ASICs 402a-408a 控制十六个千兆位以太网端口 1102-1108。端口 1102-1108 通过诸如图 8 中寄存器 714 和 718 这样的移位和保持寄存器与 ASICs 402a-408a 相耦合。总线 110 使诸如计数器，计时器，寄存器 758 和面同步逻辑 753 产生的中间面控制信号与每一个 ASICs 402a-408a 相耦合。处理器 1114 通过总线 112 控制 ASICs 402a-408a。处理器 1114 包含 CPU 模块，DRAM，FPGA 控制和以太网控制，与内存 710，控制器 712 和控制与状态寄存器 753 为本地互联板 218 提供的功能大致相同。

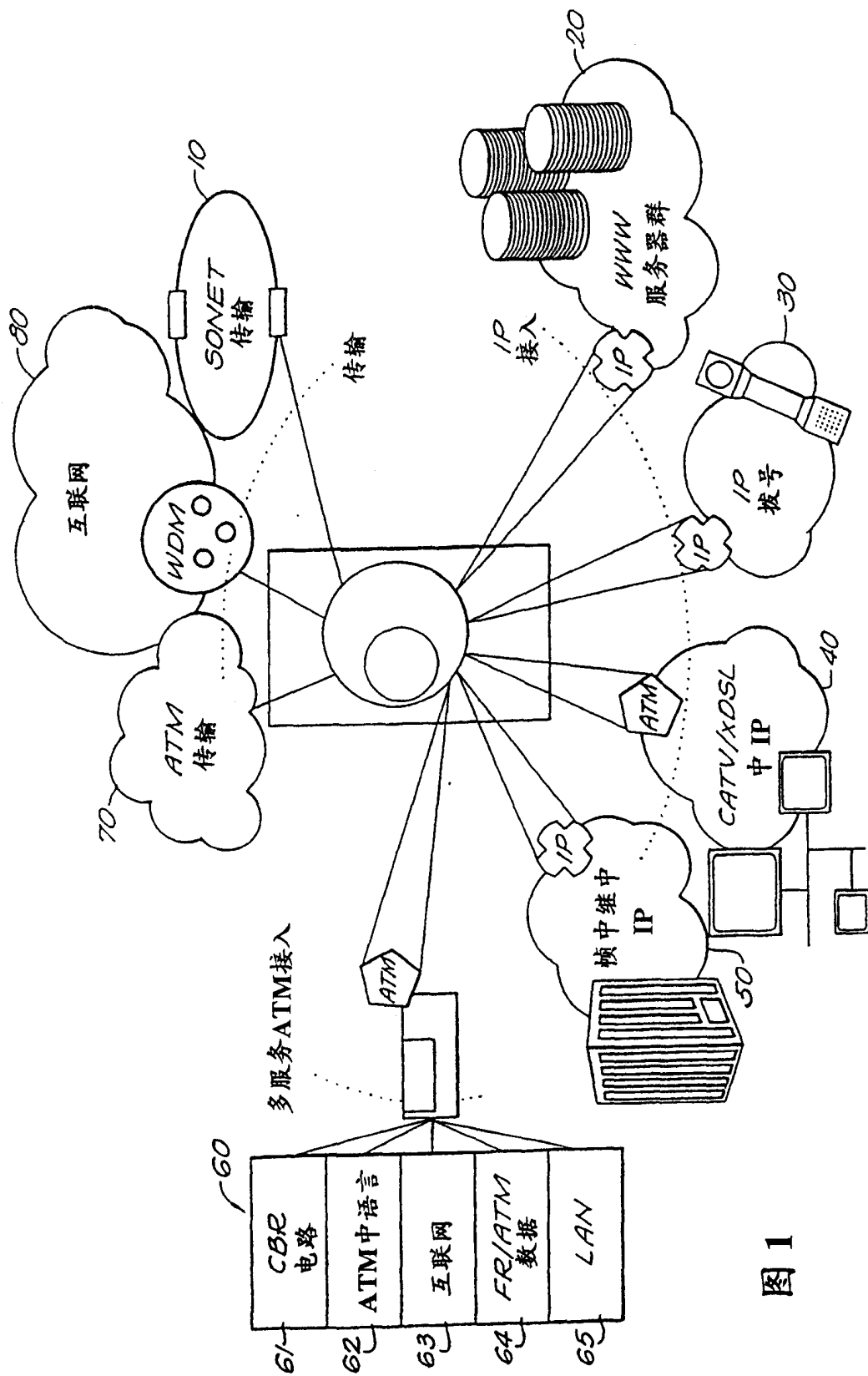
应当指出上面提到的连接和电路划分的描述可能对实际与逻辑连接和电路划分是一般来说。

图 11 说明了可以用于说明图 4 中的任何互联面的典型逻辑互联面的功能框图。

可以看到本发明实现了前面提出的目的，包含提供动态的带宽可扩展互连网络。由于上述结构和描述的方法在不背离本发明的条件下可能会有一定的改变，这意味着上面所有的与图相关的描述或说

明都是解说性的且不是被限制的。

已经说明了本发明，本发明要求保护的范围由以下权利要求书确定如下：



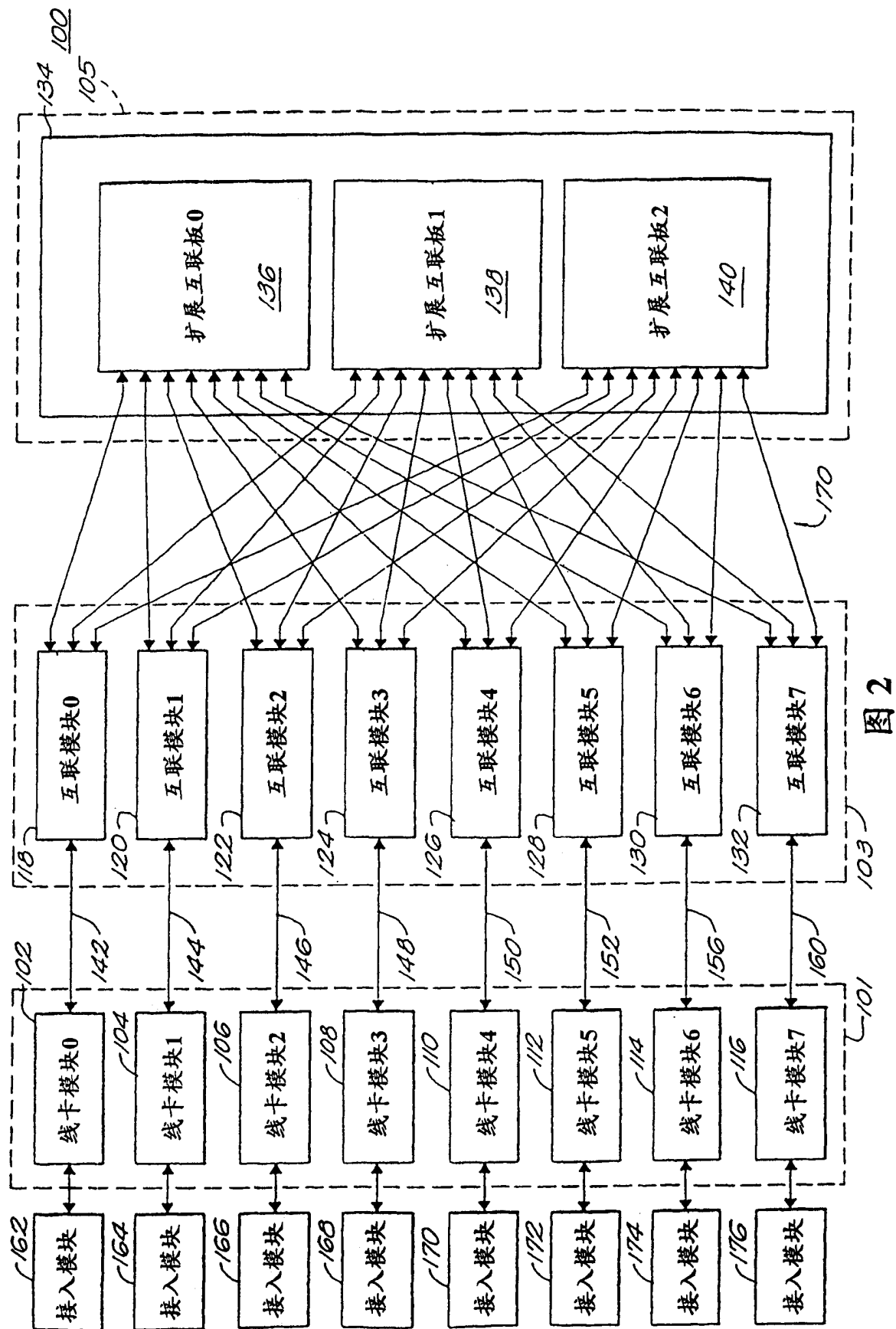


图 2

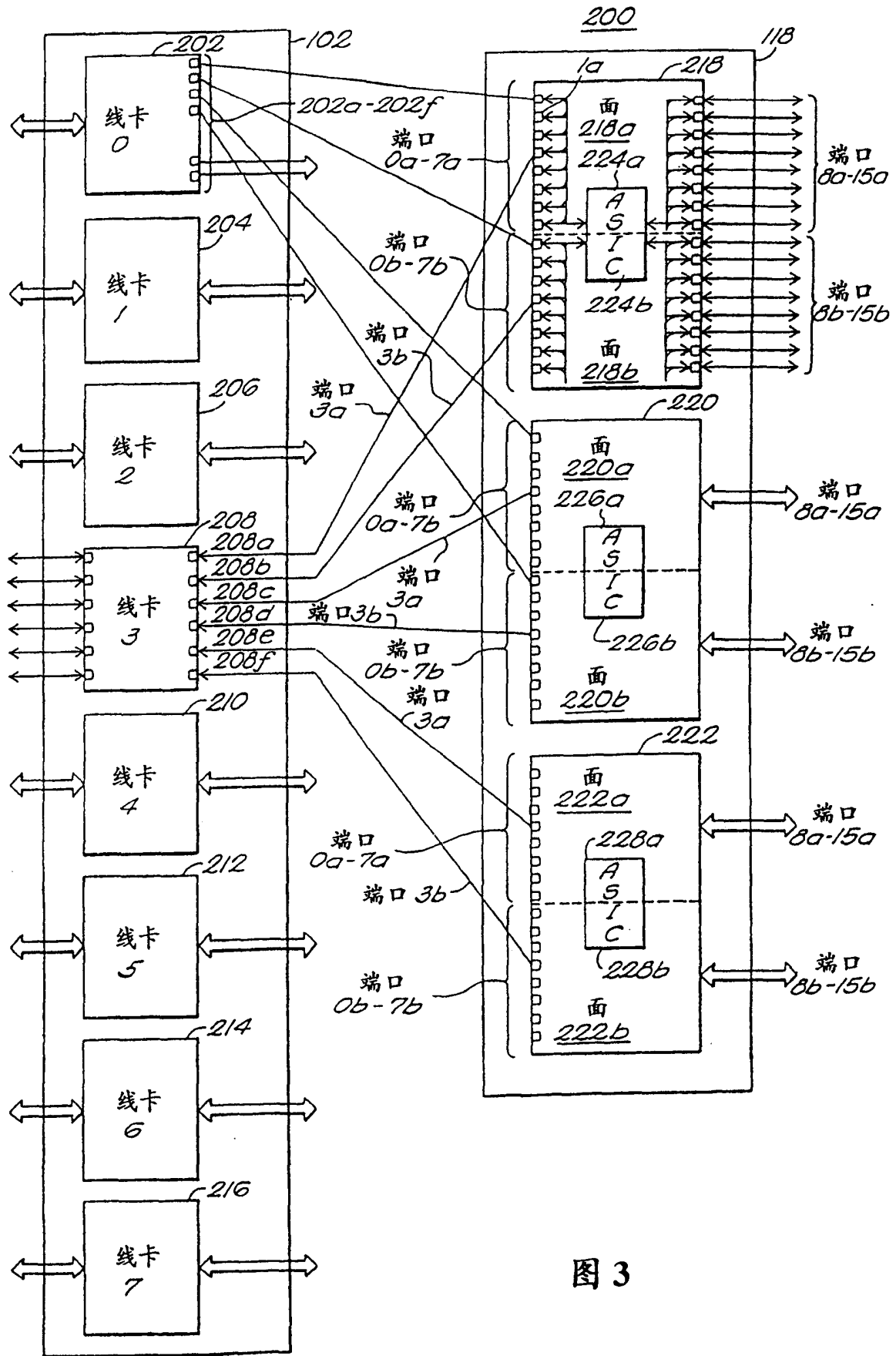


图 3

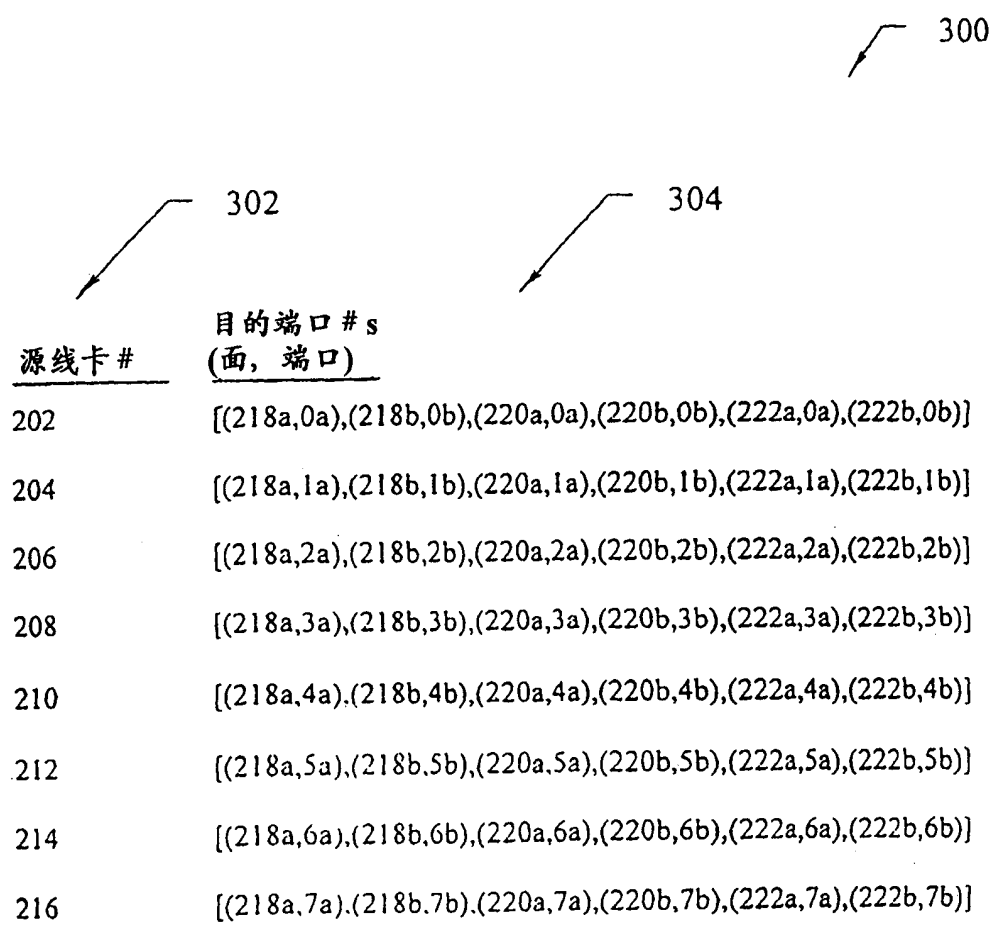


图 4

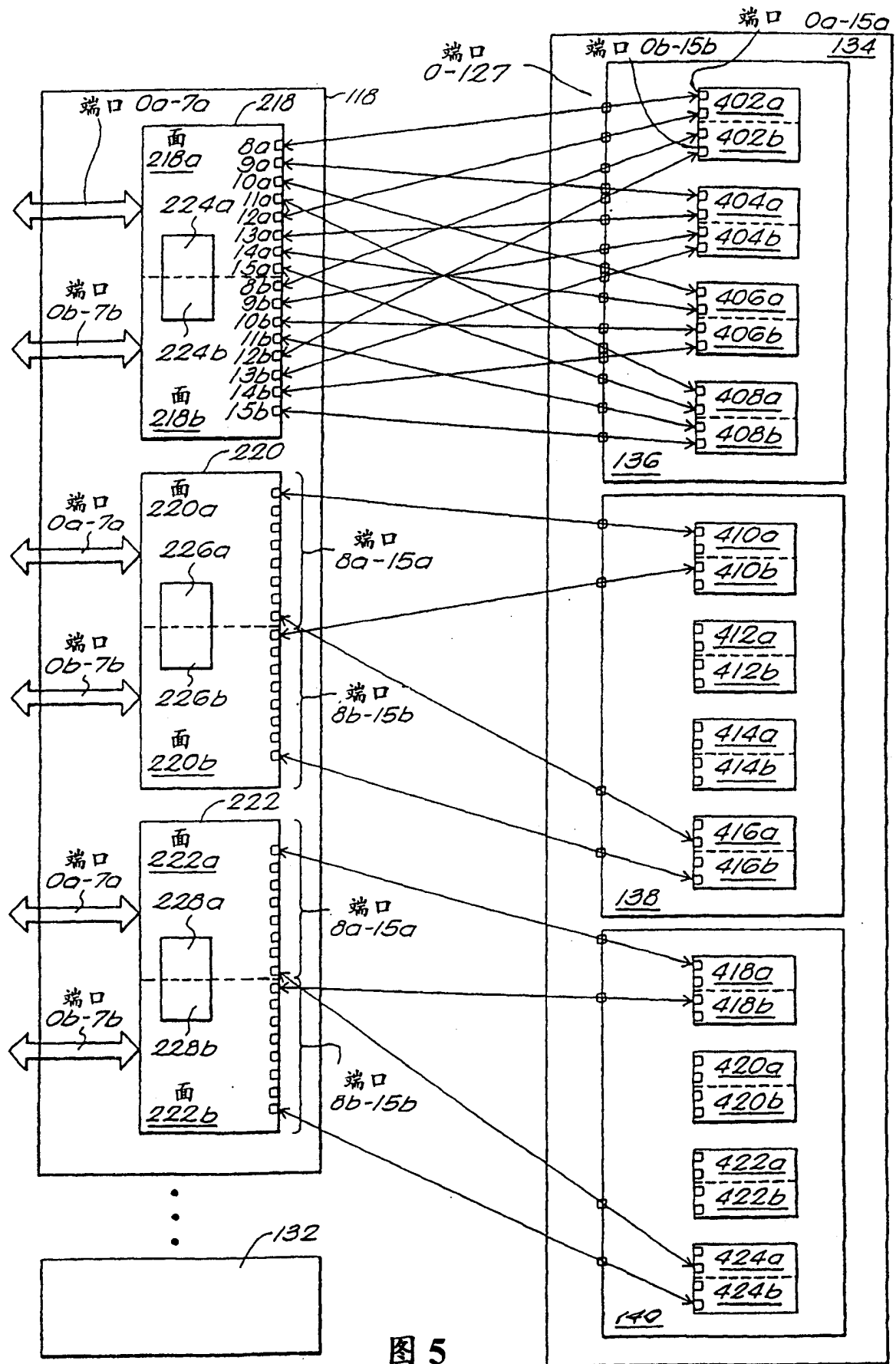


图 5

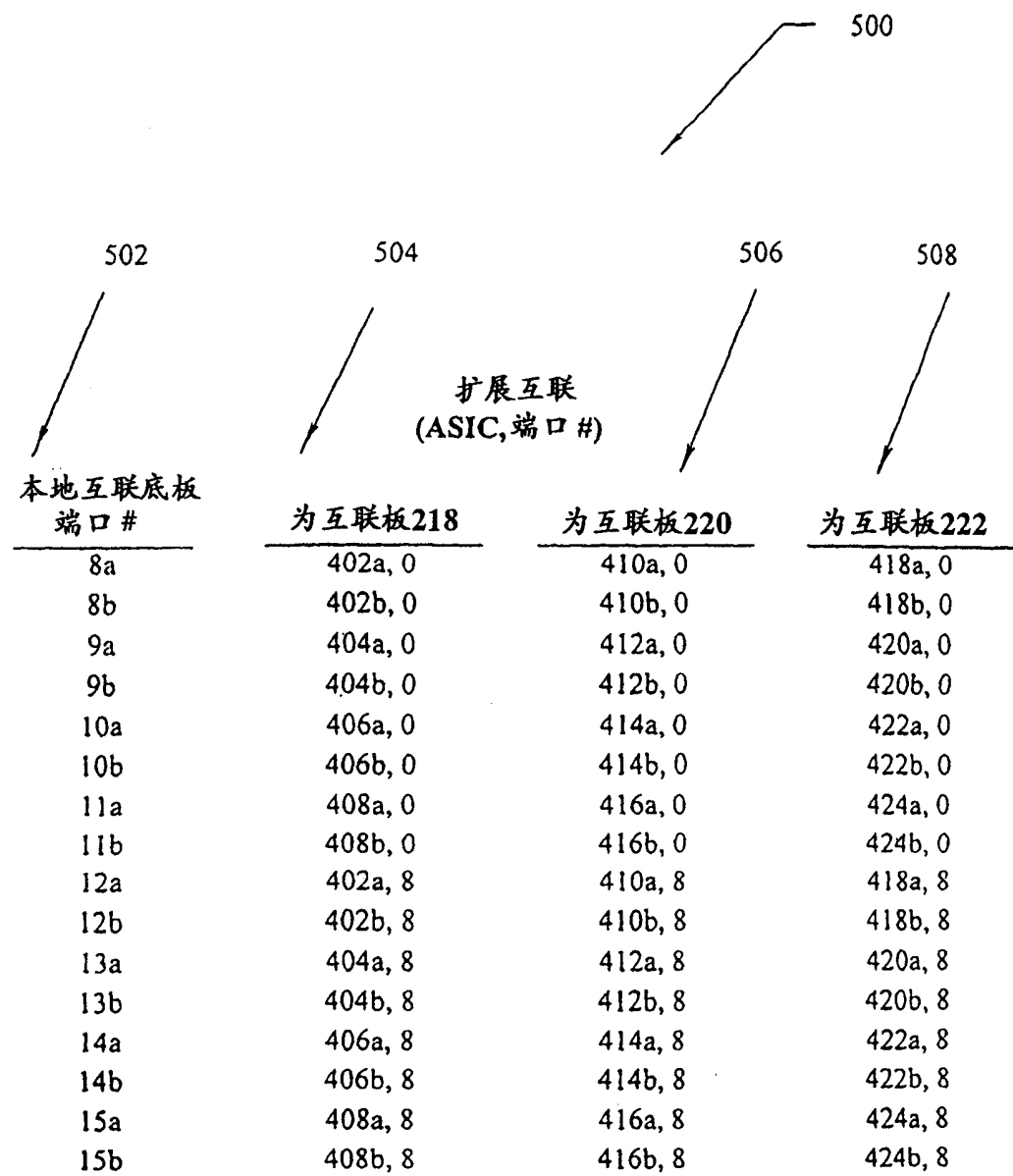


图 6

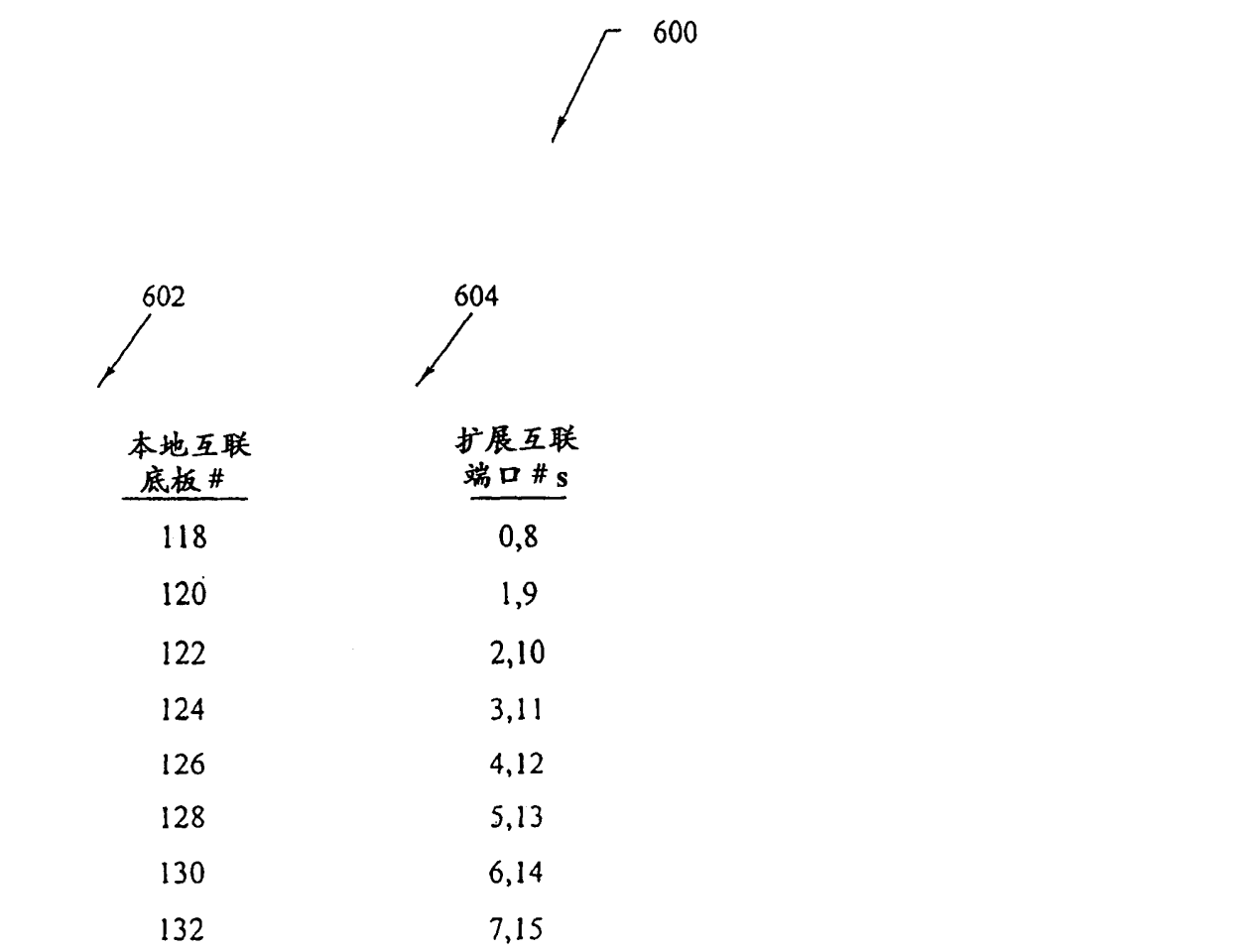


图 7

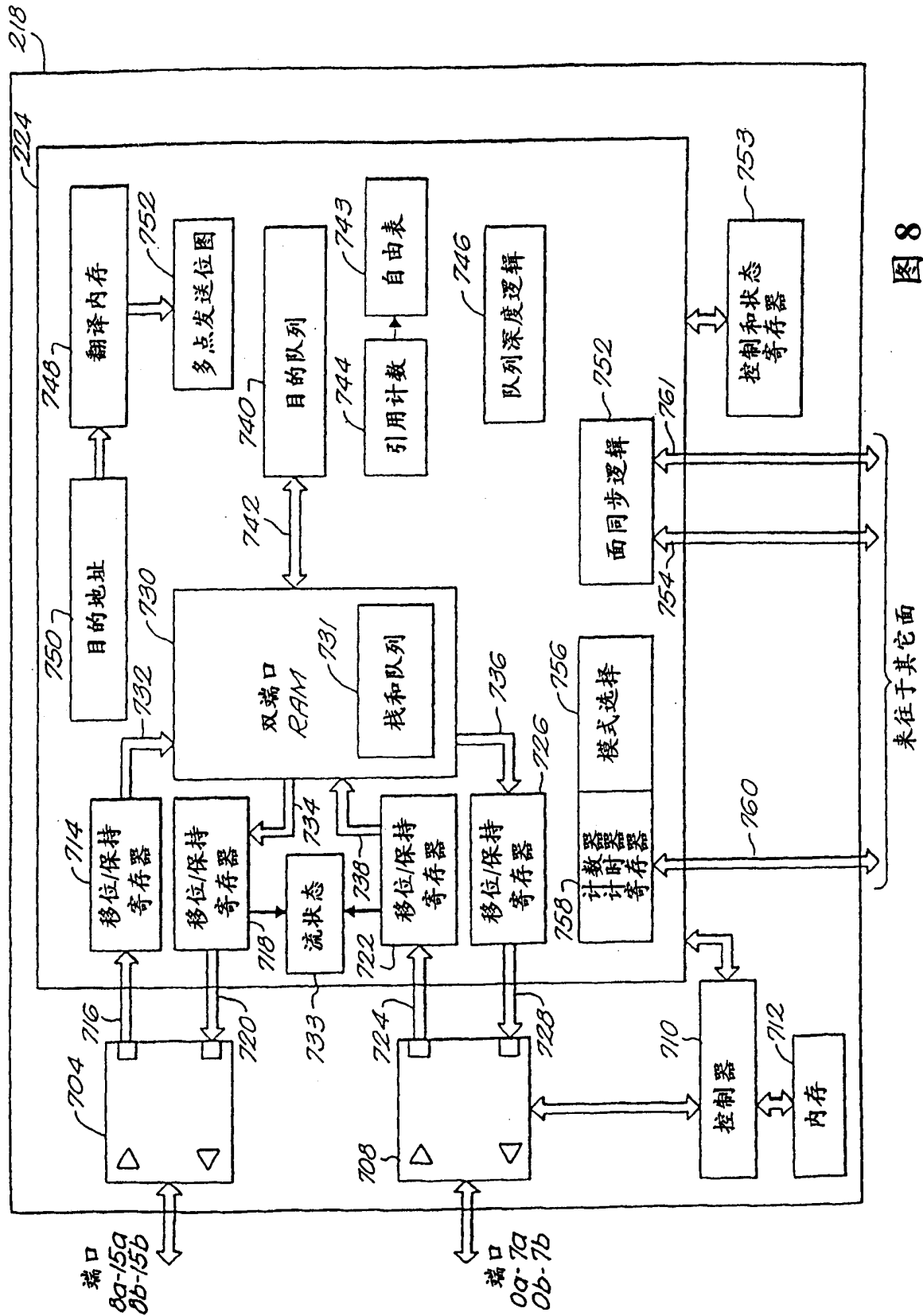


图 8

来至于其它面

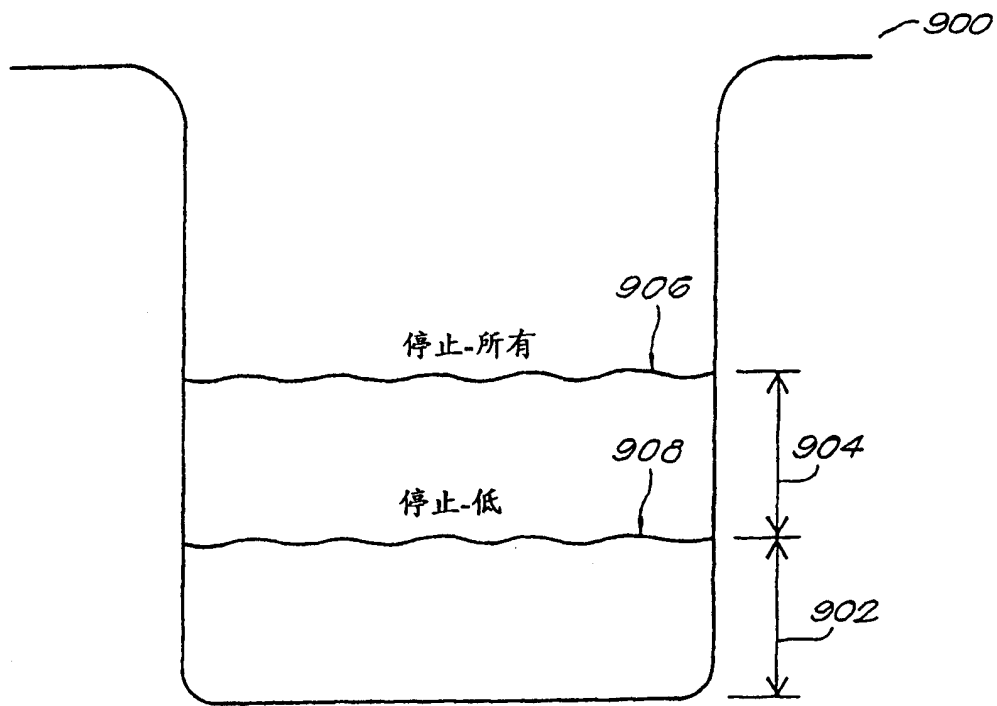


图 9A

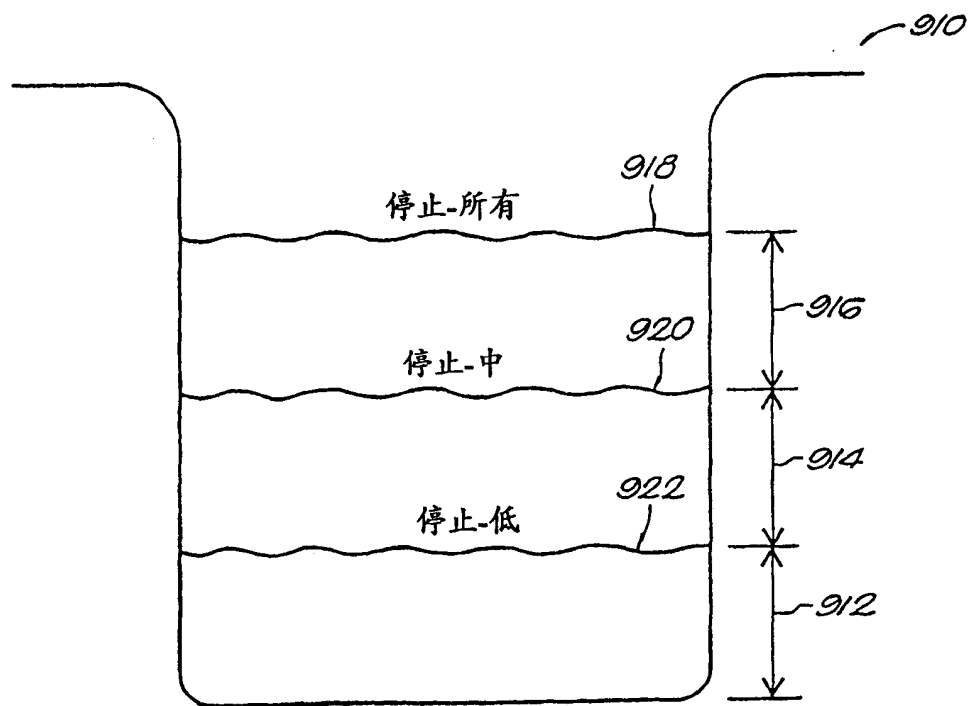


图 9B

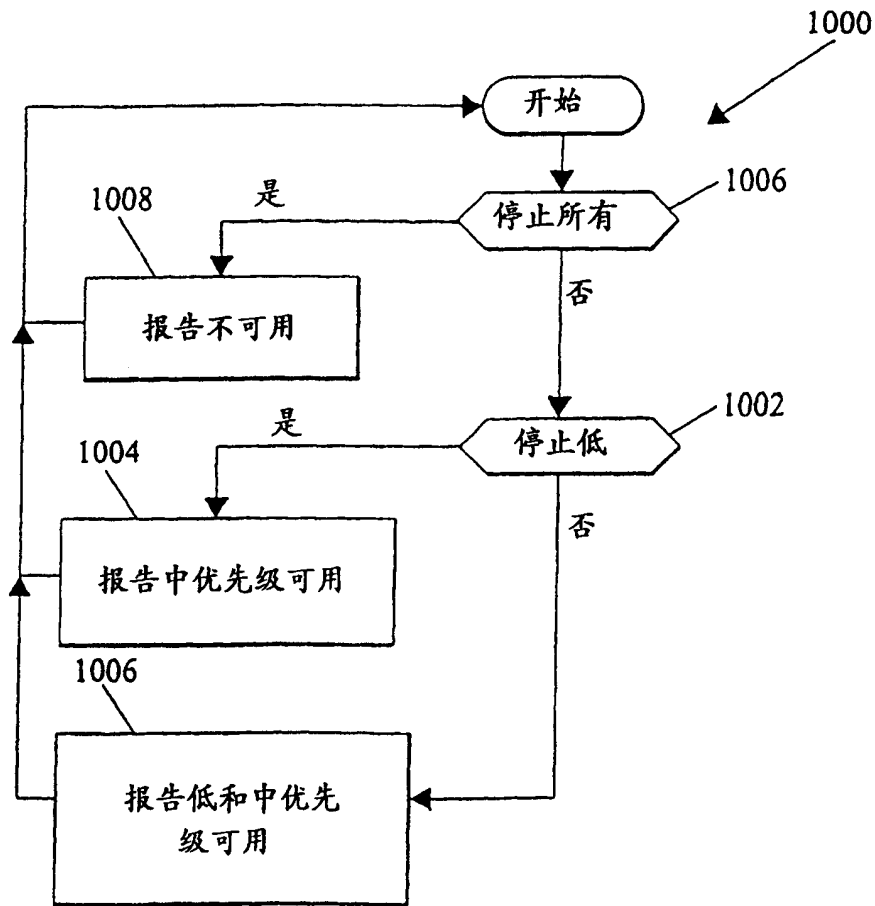


图 10A

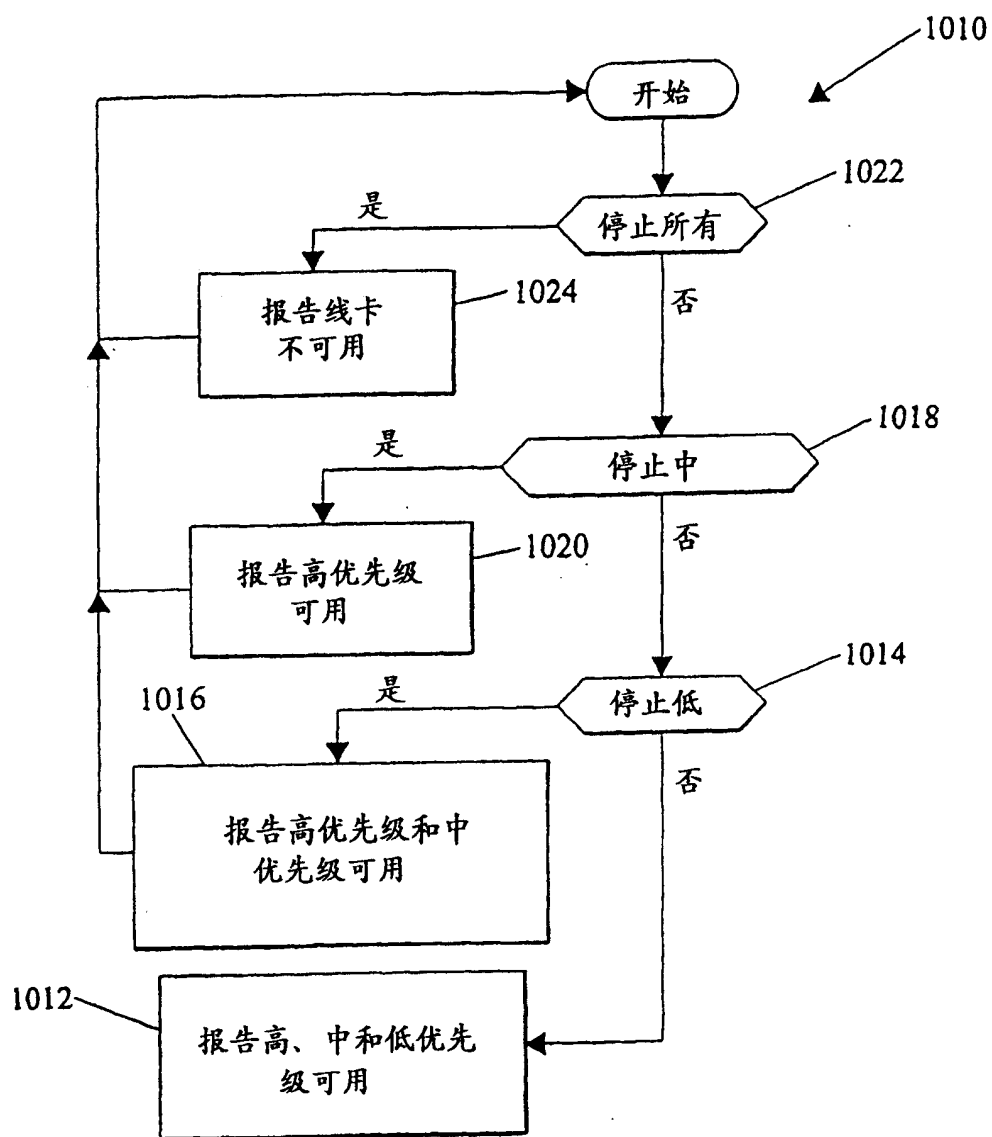


图 10B

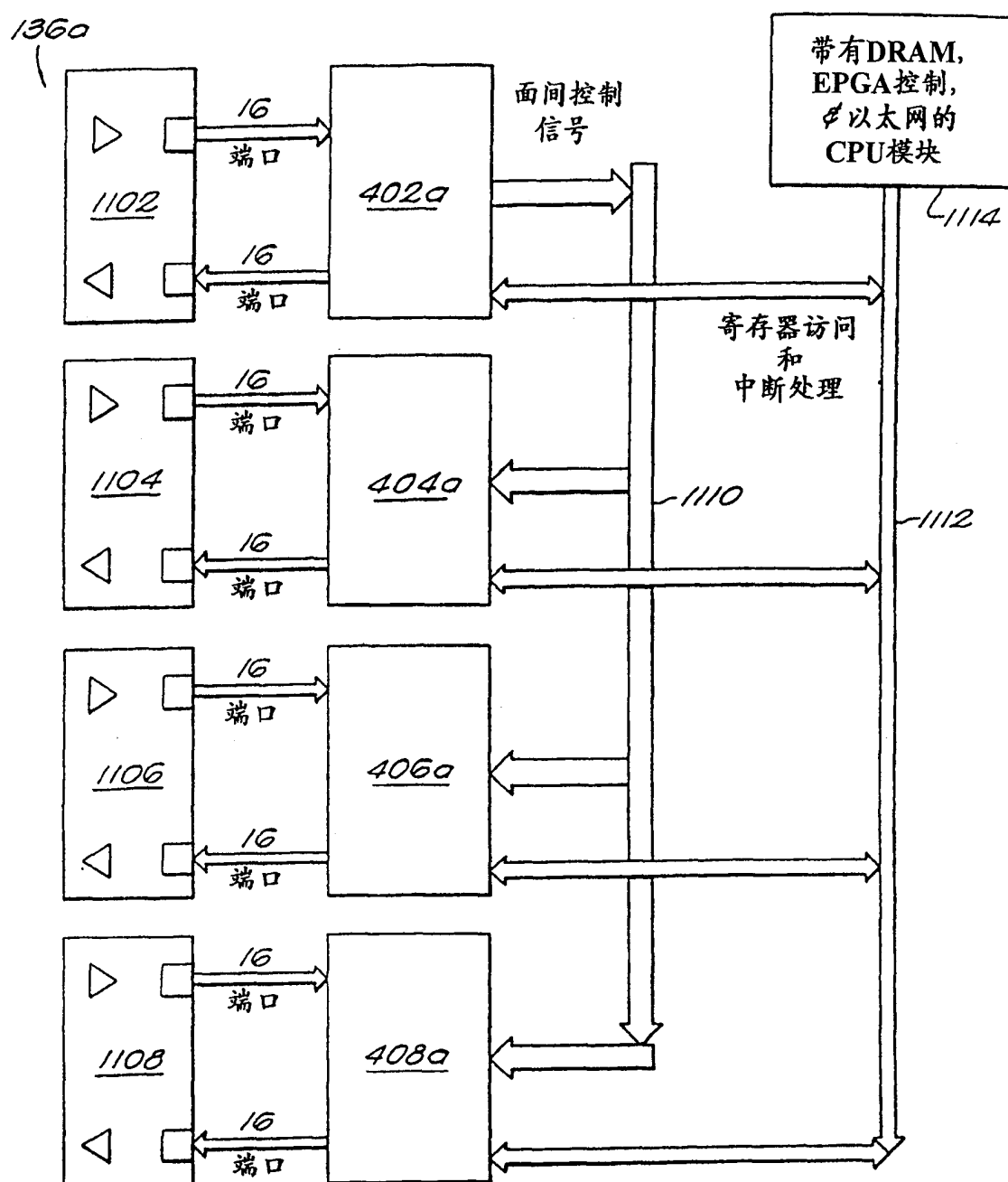


图 11