

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 13/16 (2006.01)



[12] 发明专利申请公开说明书

[21] 申请号 200480015427.3

[43] 公开日 2006年7月5日

[11] 公开号 CN 1799035A

[22] 申请日 2004.5.20

[21] 申请号 200480015427.3

[30] 优先权

[32] 2003.6.5 [33] US [31] 10/456,353

[86] 国际申请 PCT/US2004/016116 2004.5.20

[87] 国际公布 WO2004/109526 英 2004.12.16

[85] 进入国家阶段日期 2005.12.2

[71] 申请人 英特尔公司

地址 美国加利福尼亚州

[72] 发明人 P·福格特 W·莫罗

D·布尔泽津斯基

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 杨凯 王勇

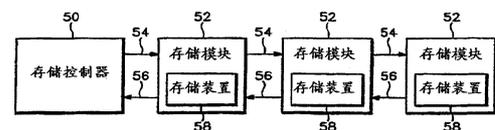
权利要求书 4 页 说明书 26 页 附图 15 页

[54] 发明名称

具有位通道故障在线恢复的存储信道

[57] 摘要

利用多个位通道的存储设备和方法可在位通道上重定向一个或多个信号。存储代理可包括具有多个位通道的再驱动电路、存储装置或接口、以及耦合在所述多个位通道与所述存储装置或接口之间的故障在线恢复电路。



1. 一种存储缓冲器，包括：
具有多个位通道的再驱动电路；
5 存储接口；以及
耦合在所述多个位通道与所述存储接口之间的故障在线恢复电路。
2. 如权利要求 1 所述的存储缓冲器，其特征在于，所述故障在线恢复电路与所述再驱动电路分开。
- 10 3. 如权利要求 1 所述的存储缓冲器，其特征在于，还包括耦合在所述多个位通道与所述故障在线恢复电路之间的偏斜消除电路。
4. 如权利要求 1 所述的存储缓冲器，其特征在于，还包括耦合在所述故障在线恢复电路与所述存储接口之间的偏斜消除电路。
5. 如权利要求 1 所述的存储缓冲器，其特征在于，所述故障在线恢复电路包括交叉开关。
15
6. 如权利要求 1 所述的存储缓冲器，其特征在于，所述故障在线恢复电路包括复用器。
7. 如权利要求 1 所述的存储缓冲器，其特征在于，所述存储缓冲器能够检测故障位通道。
- 20 8. 如权利要求 1 所述的存储缓冲器，其特征在于，所述存储缓冲器能够响应命令而标出故障位通道。
9. 一种存储模块，包括：
具有多个位通道的再驱动电路；
存储装置；以及
25 耦合在所述多个位通道电路与所述存储装置之间的故障在线恢复电路。
10. 如权利要求 9 所述的存储模块，其特征在于，所述存储模块能够检测故障位通道。

11. 如权利要求 1 所述的存储模块, 其特征在于, 所述存储模块能够响应命令而标出故障位通道。

12. 一种存储控制器, 包括:

具有多个位通道的单向链路接口; 以及

5 耦合到所述多个位通道的故障在线恢复电路。

13. 如权利要求 12 所述的存储控制器, 其特征在于, 所述存储控制器能够检测故障位通道。

14. 如权利要求 12 所述的存储控制器, 其特征在于, 所述存储控制器能够发出命令, 该命令指导代理标出所述多个位通道中有故障的一个位通道。

15. 一种存储系统, 包括:

第一代理;

第二代理; 以及

15 具有耦合在所述第一与第二代理之间的多个位通道的单向链路;

其中所述第一代理能够在所述多个位通道上重定向一个或多个信号。

16. 如权利要求 15 所述的存储系统, 其特征在于, 所述第二代理能够在所述多个位通道上重定向一个或多个信号。

17. 如权利要求 15 所述的存储系统, 其特征在于, 所述第一代理包括故障在线恢复电路。

18. 如权利要求 15 所述的存储系统, 其特征在于, 所述第一代理能够检测故障位通道。

19. 如权利要求 15 所述的存储系统, 其特征在于, 所述第一代理能够响应命令而标出故障位通道。

20. 一种用于操作存储代理的方法, 包括:

在多个位通道上再驱动信号;

将信号耦合在所述多个位通道与存储接口之间; 以及

在所述多个位通道与所述存储接口之间重定向一个或多个信号。

21. 如权利要求 20 所述的方法，其特征在于，还包括检测故障位通道。

5 22. 如权利要求 20 所述的方法，其特征在于，重定向一个或多个信号包括响应命令而重定向一个或多个信号。

23. 一种方法，包括：

在具有多个位通道的第一单向链路上从存储控制器发送信号；

10 在具有多个位通道的第二单向链路上在所述存储控制器上接收信号；以及

在所述单向链路之一中的多个位通道上重定向一个或多个信号。

24. 如权利要求 23 所述的方法，其特征在于，还包括在所述单向链路这两者中的多个位通道上重定向一个或多个信号。

15 25. 如权利要求 23 所述的方法，其特征在于，还包括检测故障位通道。

26. 如权利要求 23 所述的方法，其特征在于，还包括发出命令，该命令指导代理标出故障位通道。

27. 一种方法，包括：

20 在具有多个位通道的单向链路上从第一存储代理向第二存储代理发送信号；以及

在所述多个位通道上重定向一个或多个信号。

28. 如权利要求 27 所述的方法，其特征在于，还包括检测故障位通道。

25 29. 如权利要求 27 所述的方法，其特征在于，重定向一个或多个信号包括响应命令而重定向一个或多个信号。

30. 如权利要求 27 所述的方法，其特征在于，重定向一个或多个信号包括标出故障位通道。

31. 如权利要求 30 所述的方法，其特征在于，标出故障位通道包括在所述第一存储代理和所述第二存储代理上都重定向要发往所述故障位通道的信号。

具有位通道故障在线恢复的存储信道

5 背景

图 1 说明由电气和电子工程师学会(IEEE)作为标准提出的、非正式地称作 RamLink 的先有技术存储系统。此标准命名为 IEEE Std 1596.4-1996, 并且正式地称作基于可缩放相干接口(SCI)的高带宽存储接口信令技术(RamLink)的 IEEE 标准。图 1 的系统包括存储控制器 10 和一个或多个存储模块 12。存储控制器 10 通常内置于处理器中或者在处理器的伴随芯片组上制作。各存储模块 12 具有从接口 14, 该接口具有一个链路输入端和一个链路输出端。这些组件通过组件之间的单向链路 16 安排成称作 RingLink 的 RamLink 信令拓扑。各模块上的控制接口 18 使从接口 14 与存储装置 20 接口。在图 1 所示的系统中, 称作 SyncLink 的另一种 RamLink 信令拓扑用于从接口与存储装置之间。

RamLink 系统的目的是为处理器提供对存储装置的高速存取。数据在存储控制器与模块之间通过沿 RingLink 传播的包进行传送。控制器负责产生所有请求包以及调度从接口响应包的返回。

20 当控制器向特定模块发送包含命令、地址、时间和数据的请求包时, 发起写事务。该包从模块到模块传递, 直到它到达预期从接口, 从接口则把数据传递给存储装置之一供存储。从接口则发送响应包, 该包从模块到模块传递, 直到它到达控制器, 确认写事务完成。

25 当控制器向模块发送包含命令、地址和时间的请求包时, 发起读事务。那个模块上的从接口从存储装置之一中检索所请求数据, 并在响应包中将其返回给控制器, 响应包又从模块到模块传递, 直到它到达控制器。

图 2 说明先有技术 RamLink 从接口电路。在图 2 的电路中, 源

同步选通用于对入局数据信号定时。也就是说,伴随入局数据信号的选通信号用于对入局数据抽样。图 2 的电路采用锁相环(PLL),从分配给其它从接口电路的参考时钟信号中产生稳定的本地时钟信号。本地时钟信号用于对出局数据信号重新定时,从而避免当数据沿下游传递时的累积抖动。

附图简介

图 1 说明先有技术的 RamLink 存储系统。

图 2 说明先有技术的 RamLink 从接口电路。

图 3 说明根据本专利的发明原理的存储接口系统的一个实施例。

图 4 说明根据本专利的发明原理的存储模块的一个实施例。

图 5 说明根据本专利的发明原理的存储模块的另一个实施例以及存储缓冲器的一个实施例。

图 6 说明根据本专利的发明原理的存储系统、存储模块以及存储缓冲器的其它示例实施例。

图 7 说明根据本专利的发明原理的存储缓冲器的另一个示例实施例。

图 8 说明根据本专利的发明原理的再驱动电路的一个实施例。

图 9 说明根据本专利的发明原理的 I/O 单元的一个实施例。

图 10 说明根据本专利的发明原理的 I/O 单元的另一个实施例。

图 11 说明根据本专利的发明原理的 I/O 单元的另一个实施例。

图 12 说明根据本专利的发明原理的 I/O 单元的另一个实施例。

图 13 说明根据本专利的发明原理的故障在线恢复电路的一个实施例。

图 14 说明根据本专利的发明原理、以正常模式工作的故障在线恢复电路的另一个实施例。

图 15 说明根据本专利的发明原理、以故障在线恢复模式工作的

故障在线恢复电路的另一个实施例。

图 16 说明根据本专利的发明原理、具有位通道故障在线恢复功能的存储缓冲器的一个实施例。

5 图 17 说明根据本专利的发明原理、具有位通道故障在线恢复功能的存储控制器的一个实施例。

图 18 说明根据本专利的发明原理实现置换状态模式的一个实施例。

图 19 说明根据本专利的发明原理的置换模式发生器的一个实施例。

10 图 20 至图 23 说明根据本专利的发明原理的状态模式的示例实施例。

图 24 说明根据本专利的发明原理的存储代理的一个实施例。

图 25 说明根据本专利的发明原理的轮询操作的一个实施例。

15 详细描述

本专利包含具有独立实用性的众多发明。在一些情况下，当原理中的一些在各种相互组合中应用时，实现额外的益处，从而产生附加的发明。这些原理可在各种实施例中实现。虽然为了说明发明原理而给出一些具体细节，但是，根据本专利的发明原理，也可设计其它许多方案。因此，这些发明原理不限于本文所公开的具体细节。

20 图 3 说明根据本专利的发明原理的存储接口系统的一个实施例。图 3 的系统包括存储控制器 50 以及通过由单向链路构成的信道进行通信的一个或多个存储模块 52。信道具有包括一个或多个出站链路 54 的出站通路以及包括一个或多个进站链路 56 的进站通路。各模块可以能够在出站通路上以及在进站通路上从链路到链路再驱动信号。例如，如果模块检测到它是最外面的模块，或者响应来自存储控制器的命令，各模块还可以能够有选择地禁用任何再驱动特征。

各模块包括设置成向和/或从通路中的一个或多个传送数据的一

个或多个存储装置 58。例如，模块可经过设置，使得来自出站通路的数据被传送到存储装置，而来自存储装置的数据则被传送到进站通路。一个或多个缓冲器可设置在一个或多个存储装置与通路中的一个或多个之间。模块和控制器不限于任何特定的机械装置。例如，模块可在与系统的其它部分分开的衬底上制作，它们可在与控制器和链路公共的衬底上制作，或者它们可在其它任何机械装置中实现。模块也不限于任何特定类型的存储装置，例如只读存储器(ROM)、动态随机存取存储器(DRAM)、闪速存储器等。

图 4 说明根据本专利的发明原理的存储模块的一个实施例。图 4 的模块包括两个再驱动电路 60 和 62，分别接收单向链路 54A 和 56A 上的信号以及将信号再驱动到单向链路 54B 和 56B 上。一个或多个存储装置 58 设置成向和/或从再驱动电路中的一个或多个传送数据。

图 4 的模块不限于单向链路的任何特定布置或者向和/或从再驱动电路传送数据的任何特定布置。如果图 4 的模块要用于例如图 3 所示的存储系统，则再驱动电路 60 可被指定为出站再驱动电路，并设置成接收和再驱动包括链路 54A 和 54B 的出站通路上的信号，而另一个再驱动电路 62 则可被指定为进站再驱动电路，并设置成接收和再驱动包括链路 56A 和 56B 的进站通路上的信号。在这个实例中，一个或多个存储装置 58 可经过设置，使得数据从出站再驱动电路 60 传送到存储装置，以及从存储装置传送到进站再驱动电路 62。

模块可以能够检测它是否为信道上最外面的模块，并且相应地选择性地禁用任何再驱动特征。例如，如果图 4 的模块将用于如图 3 所示的存储系统，以及模块检测到它是最外面的模块，则出站再驱动电路接收链路 54A 上的入局信号，但不对它们再驱动。同样，进站再驱动电路仅采用对应于从存储装置接收的数据的信号和/或可由模块内部产生的其它信号来驱动链路 56B。或者，即使模块检测到它不是最外面的模块，它也可经过构建，使得它能够好像它是最外面模块一样工作(例如响应来自存储控制器的命令)，在这种情况下，它可忽略

在入站链路 56A 上接收的信号，并且它可能不把信号再驱动到出站链路 54B。

图 5 说明根据本专利的发明原理的存储模块的另一个实施例以及存储缓冲器的一个实施例。图 5 的模块包括具有两个再驱动电路 60 和 62 的存储缓冲器 64，分别接收单向链路 54A、56A 上的信号以及将信号再驱动到单向链路 54B、56B 上。存储缓冲器还包括设置成向和从一个或多个存储装置 58 传送数据的存储接口 66。缓冲器可以能够检测它是否为信道上的最后一个代理，并且相应地有选择地禁用任何再驱动特征。即使缓冲器不是最后一个代理，它也可以能够好像它是信道上的最后一个代理那样工作，例如响应来自存储控制器的命令。本文所使用的代理表示与信道接口的任何存储控制器(又称作主控制器)、模块、缓冲器等。

图 5 的模块和缓冲器不限于单向链路的任何特定布置或者用于在存储接口与再驱动电路之间传送的任何特定布置。如果图 5 的模块要用于例如图 3 所示的存储系统，则再驱动电路 60 可能被指定为出站再驱动电路，并设置成接收和再驱动包括链路 54A 和 54B 的出站通路上的信号，而另一个再驱动电路 62 则可被指定为入站再驱动电路，并设置成接收和再驱动包括链路 56A 和 56B 的入站通路上的信号。在这个实例中，存储接口可配置成从出站再驱动电路 60 接收数据以及向入站再驱动电路 62 发送数据。

各种机械布置可用来实现图 4 和图 5 的存储模块和/或缓冲器。例如，存储装置 58、再驱动电路 60 和 62、以及缓冲器 64 均可实现为安装在公共电路板或者在分开的电路板上的分开的集成电路。组件的各种组合可在公共集成电路上制作在一起，或者它们都可制作在单个集成电路上。一个或多个电路板(若有的话)可以能够插入主板上的插槽、与主板制作在一起、或者以另外任何方式来设置。例如，如果组件作为多芯片模块的组成部分来制作，则可能没有电路板。根据本专利的发明原理的存储缓冲器可用于将不同于存储装置的装置与信

道接口。例如，根据本专利的发明原理的存储缓冲器可用于将 I/O 控制器或桥与信道接口。

5 参照“进站”和“出站”通路、链路、再驱动电路等来描述根据本专利的发明原理的设备的其它实施例，以帮助理解该设备可如何用于例如图 3 所示的实施例的存储系统。但是，这些设备不限于单向链路的任何特定布置、用于在链路与其它电路之间传送数据所示的特定布置或者所示的实现详细情况的任一个。

10 图 6 说明根据本专利的发明原理的存储系统、存储模块以及存储缓冲器的其它示例实施例。参照图 6，一个或多个存储模块 52 基于具有沿一侧面的两边的接触片的印刷电路板，以便创建可插入固定系统的其它元件的另一个电路板上的连接器中的双列直插式存储模块 (DIMM)。现有形状因数可用于此模块，例如用于双数据速率 II(DDR2) 动态随机存取存储器(DRAM)模块的 DIMM 形状因数。

15 模块装载了存储装置 58，例如商品型 DRAM、如 DDR2 DRAM。各模块上的存储缓冲器 64 隔离存储装置与将模块和存储控制器 50 接口的信道，存储控制器 50 又称作主控制器。信道以点到点布置与包括出站链路 54 的出站通路连接以及与包括进站链路 56 的进站通路连接。链路可采用低电压差动信号通过并行单向位通道来实现。

20 在图 6 的实施例中，没有附加信号线用于诸如命令、重置、初始化等功能。而是，这些功能直接在通过信道发送的数据中编码。然而，作为选择，任何数量的附加信号线可用来实现这类功能。

25 参考时钟信号 REF CLK 由时钟合成器 76 来产生，也许通过时钟缓冲器 78 分配给主控制器和模块。这有助于准异步定时方案，其中，本地产生的时钟信号用于对入局数据抽样和再驱动。由于公共参考时钟在各代理处可得到，因此数据信号可被定时而没有任何频率跟踪。或者，本地时钟信号可与任何参考时钟无关地产生。作为另一个备选方案，可采用同步定时方案、如源同步选通。

在一个可能的实施例中，主控制器通过向出站通路上最里面的模

块-也许采用包或帧(这些术语在这里可互换使用)-发送数据来发起数据传递。最里面的模块接收该数据并再驱动到出站通路上的下一个模块。各模块接收并再驱动出站数据,直到出站数据到达最外面的模块。虽然最外面的模块可尝试把数据再驱动到“不存在的”出站链路,但各模块可以能够检测(或被指示)它是最外面的模块,并禁用任何再驱动电路以降低不必要的功耗、噪声等。在这个实施例中,朝主控制器的方向、即入站的数据传递由最外面的模块发起。各模块接收入站数据并将它沿入站通路再驱动,直到它到达主控制器。

在物理信道上可使用任何适当的通信协议。例如,可指定主控制器来发起和调度所有入站及出站数据传递。或者,可允许任何代理来发起数据传递。数据帧可配置成携带命令、读数据、写数据、状态信息、错误信息、初始化数据、空闲模式等或者它们的任何组合。协议可实现成,当主控制器沿出站通路向目标模块发送命令帧时,目标模块通过立即沿入站通路向主控制器发回响应帧来进行响应。在这种实施例中,目标模块不在出站通路上再驱动命令帧。

在一个备选实施例中,目标模块接收命令帧,然后在出站通路上再驱动命令帧。当最外面的模块接收命令帧时,它在入站通路上发起响应帧(也许只是空闲帧)。目标模块等待,直到响应帧到达其入站接收器。然后目标模块例如通过用目标模块的真实响应帧取代最外面的模块所发送的响应帧,将其响应合并到入站数据流中。

图 7 说明根据本专利的发明原理的存储缓冲器的另一个示例实施例。图 7 的存储缓冲器包括接收和再驱动包括链路 54A、54B 的出站通路上的信号的出站再驱动电路 60 以及接收和再驱动包括链路 56A、56B 的入站通路上的信号的入站再驱动电路 62。存储接口 66 将缓冲器与一个或多个存储装置接口,这可通过存储总线 68 进行。存储接口可包括读和/或写缓冲器、如 FIFO 缓冲器。来自出站通路的数据耦合到存储接口,这可通过消除在出站通路具有不止一个位通道时数据位之间的偏斜的偏斜消除电路 70 来进行。模式发生器 72 可用

于例如在缓冲器正好是信道上最外面的代理时产生要传送到进站通路上的状态模式，在这种情况下，在入局进站链路 56A 上可能没有接收到信号。复用器 74 有选择地把来自存储接口或模式发生器的数据耦合到进站再驱动电路。

5 存储接口不限于任何特定布置，并且它可与标准存储装置、尤其是商品存储装置、如 DDR2 DRAM 兼容。整个存储缓冲器可在单个集成电路上集成，可集成到一个或多个存储装置中，其组成元件可集成到分开的组件中，或者可采用其它任何机械布置。图 7 所示的实施例仅作为示范，根据本专利的发明原理，其它实施例是可行的。例如，
10 图 7 的实施例采用从出站再驱动电路流动到存储接口以及从存储接口流动到进站再驱动电路的单向数据来表示。但是，这种数据流可以是双向的，并且考虑其它布置。即使图 7 的实施例要用于其中存储接口的数据只需要按图 7 所示流动的信道系统中，它也仍然可采用具有全双向数据访问的再驱动电路来实现，因为这可有助于例如内置自检
15 (BIST)功能的实现，在这种情况下，用于对来自进站通路的数据进行偏斜消除的第二偏斜消除电路可能会有帮助。

 图 8 说明根据本专利的发明原理的再驱动电路的一个实施例。图 8 的电路包括一个或多个输入/输出(I/O)单元 74，其中的每个接收可再驱动为输出数据信号 TX 的输入数据信号 RX。或者，I/O 单元可将
20 读数据信号 RDX 替换或合并到输出数据信号中。写数据信号 WDX 可在输入数据信号被再驱动为输出数据信号之前或者之后从其中提取。

 上述信号名称的任一个中的“X”表明它可能是取决于再驱动电路中的 I/O 单元的数量多个相似信号其中之一。例如，具有九个位
25 通道的再驱动电路将包括具有命名为 R0、R1...R8 的输入数据信号的九个 I/O 单元。在仅具有单个 I/O 单元的再驱动电路中，数据输入信号是 R0 或只是 R。术语 RX 用于一般表示输入数据信号的任一个或全部。

为了方便起见，术语“写数据”用来表示从通过 I/O 单元传播的数据流中提取的任何数据。但是，这并不意味着写数据必须针对存储接口或存储装置。同样，“读数据”表示输入到 I/O 单元的任何数据，但读数据可来自任何源，而不只是存储装置或存储接口。

5 再参照图 8，时钟发生器 80 响应参考时钟信号 REF CLK 而产生多个相位时钟信号 PCX 和发送时钟信号 TC。时钟发生器包括产生作为参考时钟信号 REF CLK 的倍数的发送时钟 TC 的锁相环(PLL)82 以及相位时钟发生器 84。在一个可能的实施例中，存在间隔 90 度且从发送时钟 TC 得到的四个相位时钟信号 PC0、PC1、PC2 和 PC3。I/O
10 单元中的每个可采用 TC 和 PCX 时钟信号中的一个或多个来对数据信号抽样和/或再驱动，和/或产生附加的本地时钟信号。在这个实施例中，相位时钟和发送时钟信号在它们不随输入数据信号 RX 中任一个的相位而调整的意义上是稳定信号。

图 9 说明根据本专利的发明原理的 I/O 单元的一个实施例。接收器 86 设置成接收数据信号 RX，并响应抽样时钟信号 SC 而将它再驱动为数据信号 TX。抽样时钟信号由抽样时钟发生器 88 产生，抽样时钟发生器 88 能够响应数据信号 RX 而调整抽样时钟信号。写数据信号 WDX 可从接收器 86 的输入或输出中提取。如果如图 9 所示从接收器的输出中提取，则抽样时钟信号 SC 可用作或者用于导出写数据的选通信号。对抽样时钟发生器的输入可从与图 9 所示的接收器的输入不同的点提取。例如，它也可从接收器的输出中提取。

图 10 说明根据本专利的发明原理的 I/O 单元的另一个实施例。在图 10 的实施例中，抽样时钟发生器 88 采用内插器 90 和接收器跟踪单元(RTU)92 来实现。内插器通过响应来自接收器跟踪单元的跟踪信号而在多个相位时钟信号 PCX(在本例中为 90 度异相的四个信号)之间内插来产生抽样时钟信号。接收器跟踪单元观察数据信号 RX 并调整跟踪信号，使得抽样时钟信号使接收器在适当时间对数据信号进行抽样和再驱动。这样，抽样时钟信号可动态跟踪数据信号。

在一个可能的实施例中，接收器跟踪单元通过对数据信号过抽样并调整抽样时钟信号来观察数据信号 RX 中的转变，从而在数据眼图的中心、即在数据信号中转变之间的中间点对数据信号进行抽样和再驱动。抽样时钟发生器 88 可包括环路滤波器，它测量若干位单元，
5 并且可最终确定它应当调整抽样时钟信号的相位以捕捉更接近数据眼图位置的中心的的数据。对抽样时钟发生器的输入可从与图 10 所示的接收器的输入不同的点提取。例如，它也可从接收器的输出中提取。

根据本专利的发明原理的 I/O 单元的一个实施例可与训练 I/O 单元以动态跟踪数据信号的方案配合使用。例如，如果图 10 的 I/O 单元用作图 3 所示的存储模块之一，则主控制器可定期把训练帧发送到出站通路上。这些训练帧的边沿密度足以保证接收器跟踪单元观察数据信号中的足够转变，从而能够调整抽样时钟信号。同样，图 3 中最外面的模块可定期把训练帧发送到进站通路上。

图 11 说明根据本专利的发明原理的 I/O 单元的另一个实施例。图 11 的实施例与图 9 类似，只是在数据信号通路中增加了缓冲器 94。缓冲器 94 可以是补偿电压和温度引发效应的防抖动或漂移补偿缓冲器。缓冲器使数据信号 TX 与发送时钟信号 TC 再同步。发送时钟信号在其相位不象抽样时钟信号 SC 那样随数据信号而调整的意义
15 上是稳定的。

在图 11 的实施例中，缓冲器能够响应模式信号而以通过模式或者再驱动模式工作。在通过模式中，信号没有被抽样和再驱动就通过。在再驱动模式中，信号响应时钟信号而被抽样和再驱动。这使 I/O 单元能够以不同的再驱动模式工作。在一个可能的实施例中，缓冲器
25 在模式信号被断言时以通过模式工作。这称作再抽样模式，并且可产生更短的等待时间，因为数据信号由用于对数据抽样的相同时钟来再驱动。当模式信号未被断言时，缓冲器以再驱动模式工作，因此数据与发送时钟再同步。这称作再同步模式，并且可产生更长的等待时间

但可减小抖动。I/O 单元可设计到具有用于接收模式信号的输入端的存储缓冲器或模块中。如果存储缓冲器或模块要在其中到下一个代理存在较短信号通路的系统中使用，则输入可被断言(或不被断言，取决于极性)，从而使 I/O 单元以再抽样模式工作，因为在短信号通路上可容许更大抖动。另一方面，如果存储缓冲器或模块要在其中到下一个代理存在较长信号通路的系统中使用，则输入可被取消断言，以便使 I/O 单元以再同步模式工作，因为这减小抖动，虽然可能的代价是较长的等待时间。或者，已登记标志可用于存储缓冲器或模块、或者再驱动电路或 I/O 单元本身中来控制模式信号。

10 图 12 说明根据本专利的发明原理的 I/O 单元的另一个实施例。在图 12 的实施例中，接收和发送数据信号 RX 及 TX 为差动信号，并表示为穿过在其中可制作 I/O 单元的集成电路管芯的一边。接收器 86 包括抽样单元 96 和端接单元 98。抽样单元响应由内插器 90 响应来自抽样时钟发生器的相位时钟信号而产生的抽样时钟信号 SC 而对入局数据信号抽样。端接单元提供差动端接，并把差动数据信号转换为单端信号。防抖动或漂移补偿缓冲器 94 响应抽样时钟信号 SC 或稳定的发送时钟信号 TC 而对数据定时。复用器 100 有选择地将来自缓冲器 94 或串行化器 102 的数据信号耦合到发送锁存器 104。读数据信号 RDX[0...n]在串行化器 102 上的 I/O 单元被接收。另一个复用器可设置在缓冲器 94 与发送锁存器 104 之间，发送锁存器 104 的一个输入端连接到缓冲器，另一个输入端连接到内插器的输出端。

25 当 I/O 单元需要将读数据合并到数据流时，复用器选择其耦合到串行化器的输入端，使得发送锁存器响应发送时钟信号 TC 而对来自 I/O 单元的读数据定时。否则，复用器选择来自缓冲器的数据信号，该数据信号然后由发送锁存器再驱动。发送数据信号在被驱动到下一个单向链路上之前由发送器 106 重新转换为差动信号。写信号从发送锁存器的输出中提取，在解串器 108 中被收集，然后被路由到偏斜消除电路、位通道故障在线恢复装置或其它电路。解串器还可提供位线

时钟信号 BLC，它可从抽样时钟信号中得到，以便表明写数据 WDX[0...n]何时是有效的。

5 本专利的发明原理的一部分涉及与再驱动通路分离的偏斜消除信号。再驱动通路由信号在被接收和再驱动时传播所通过的一个或多个组件来定义。例如，在图 9 和图 10 的实施例中，再驱动通路包括接收器 86。在图 11 的实施例中，再驱动通路包括接收器 86 和缓冲器 94。在图 12 的实施例中，再驱动通路包括抽样单元 96、端接单元 98、缓冲器 94、复用器 100、发送锁存器 104 以及发送器 106。

10 根据本专利的发明原理的一部分，偏斜消除电路可集成到再驱动电路中，使得偏斜消除电路的各个位通道包含在再驱动通路中。因此，位通道上的信号可在沿通路被再驱动时在每个再驱动电路中消除偏斜。但是，作为选择，根据本专利的发明原理的偏斜消除电路可与再驱动通路分离。例如，在图 7 的实施例中，偏斜消除电路表示为不仅与再驱动电路 60 中的再驱动通路分离，而且与整个再驱动电路分
15 离。或者，根据本专利的发明原理的偏斜消除电路可集成到再驱动电路中，但仍然与再驱动通路分离。例如，在图 12 的实施例中，一个或多个偏斜消除锁存器可设置在串行化器 102 的输出端和/或解串器 108 的输入端。

20 如上所述用于对与再驱动通路分离的信号进行偏斜消除的方法及设备的实施例仅作为示范，并不限于这些具体实例。此外，根据本专利、涉及对与再驱动通路分离的信号进行偏斜消除的原理与本专利的其它发明原理无关。例如，正如图 9-12 所示的再驱动电路的实施例不限于用于具有分开的出站和进站通路的存储系统中那样，根据本专利、涉及用于对与再驱动通路分离的信号进行偏斜消除的原理也可
25 与利用单向链路的其它类型的存储体系结构配合使用，例如利用链路的环型布置、如 RamLink 的体系结构。

本专利的发明原理的一部分涉及处理故障位通道。例如，图 3、4、5、6 或 7 的实施例中所示的任何代理之间的单向链路的任一个可

具有一个以上位通道。根据本专利的发明原理，一个或多个信号可在位通道上重定向，以便避免不良位通道。诸如存储控制器(主控制器)、模块、缓冲器等任何代理可以能够在多个位通道上重定向一个或多个信号。信号可在链路的任一端或两端重定向。任何代理可以能够自动地或者通过另一个代理的协助来检测故障位通道，以及任何代理可以能够响应来自另一个代理的命令而重定向信号。

图 13 说明根据本专利的发明原理的故障在线恢复电路的一个实施例。仅为了说明，图 13 的故障在线恢复电路 110 表示为与再驱动电路 112 的实施例一起，但发明原理不限于与任何特定再驱动电路配合使用，故障在线恢复电路也不限于图 13 所示的具体细节。再驱动电路 112 包括设置成接收和再驱动单向链路上的信号的多个位通道。每个位通道被实现为具有接收器 116 和发送器 118 的 I/O 单元 114。

故障在线恢复电路表示能够重定向送往或来自多个位通道的一个或多个信号的电路。在图 13 的实施例中，故障在线恢复电路实现为具有一个或多个复用器开关 120 的复用器。每个开关具有耦合到一个位通道的第一输入端以及耦合到相邻位通道的第二输入端，使得它可将来自任一个位通道的信号重定向到其输出端。图 13 所示的实施例表示为具有六个开关，为六个位通道提供服务，但是可使用任何数量的开关和位通道，以及开关可设置成不同于所示的相邻位通道配置的各种配置。

在正常操作模式中，开关的每个把来自其第一输入端的信号送往其输出端，如图 14 所示，使得写数据信号 WD0、WD1、WD2、WD3、WD4 和 WD5 分别送往输出端 OUT0、OUT1、OUT2、OUT3、OUT4 和 OUT5。在这种实施例中，位通道之一、例如与 WD5 对应的位通道可用于对其它位通道上的数据进行检错。

如果检测到不良位通道，则复用器可按照故障在线恢复模式工作，其中操作开关中的一个或多个以标出不良位通道。例如，如果与

WD3 关联的位通道没有正确工作，则复用器开关可分别将写数据信号 WD4 和 WD5 重定向到输出端 OUT3 和 OUT4，如图 15 所示。在这种模式中，一个位通道的信号容量丢失。如果指定了位通道之一用于检错，则原本要发往不良位通道的信号可通过检错通道重新路由，以及检错功能可被禁用。

故障在线恢复电路的输出可耦合到存储接口、存储装置或者其它电路。在图 13 的实施例中，故障在线恢复电路表示为与再驱动电路分离，但它也可集成到再驱动电路中。根据本专利的发明原理的故障在线恢复电路可采用如图所示的简单复用器来实现，但其它布置、如全交叉开关也是可行的。

图 13 所示的故障在线恢复电路的实施例设置成将来自位通道的写数据耦合到其输出端。或者，根据本专利的发明原理的故障在线恢复电路的一个实施例可设置成朝相反方向传递数据，在这种情况下，输出端 OUTX 将成为接收读数据的输入端，复用器开关可称作分用器开关，以及 I/O 单元的每个可在接收器与发送器之间具有复用器，将来自故障在线恢复电路的读数据合并到位通道中。这样，复用器表示复用器和分用器。作为另一个备选方案，根据本专利的发明原理的故障在线恢复电路的一个实施例可设置成用于位通道与存储装置、存储接口或其它电路之间的双向数据流。

根据本专利的发明原理，存储缓冲器、存储模块、存储控制器(主控制器)或具有位通道故障在线恢复功能的其它代理也可具有用于检测故障位通道、重定向信号、标出不良位通道等的各种功能。例如，具有图 13 所示的故障在线恢复电路的实施例的代理可经过设计，使得它可例如通过观察另一个代理发送的测试数据模式并重定向信号以标出故障位通道，来检测故障位通道。或者，代理可经过设计，使得它可响应来自另一个代理、例如指示存储信道上的一个或多个代理的存储控制器的命令而标出故障位通道。或者，该代理可具有两种功能。

图 16 说明根据本专利的发明原理、具有位通道故障在线恢复功能的存储缓冲器的一个实施例。图 16 的实施例与图 7 相似，但还包括耦合在偏斜消除电路 70 与存储接口 66 之间的故障在线恢复电路 122。备选实施例是可行的。例如，故障在线恢复电路可设置在再驱动电路 60 与偏斜消除电路之间，或者它可集成到再驱动电路中。图 16 的实施例还包括另一个故障在线恢复电路 124，它表示为耦合在复用器 74 与再驱动电路 62 之间，但它也可集成到再驱动电路中，或者以其它方式设置。图 16 的存储缓冲器或者也可体现为存储模块，在这种情况下，存储接口被存储装置取代。

图 17 说明根据本专利的发明原理、具有位通道故障在线恢复功能的存储控制器的一个实施例。图 17 的控制器包括出站和进站单向链路接口 126、128，它们具有在这个实施例中分别包括多个发送器和多个接收器的多个位通道。故障在线恢复电路 130 和 132 分别耦合到接口 126、128 中的位通道。在图 17 的实施例中，故障在线恢复电路表示为与链路接口分离，但它们也可与这些接口结合。控制器可以能够检测故障位通道，在这种情况下，故障在线恢复电路可标出故障位通道。作为补充或替代，控制器可以能够发送指导代理标出故障位通道的命令。

现在将在包括根据本专利的发明原理的存储控制器(主控制器)、存储模块和存储缓冲器的其它实施例的完整存储信道系统的一个示范实施例的上下文中描述根据本专利的发明原理的另外的故障在线恢复方法及设备。但是，没有任何组件局限于这个示范系统或者本文所述的任何细节。

该示范系统包括具有诸如参照图 17 所述的那些功能的故障在线恢复功能的主控制器的一个实施例以及具有具备诸如参照图 16 所述的那些功能的故障在线恢复功能的缓冲器的一个或多个存储模块的实施例。在这个实例中，主控制器和模块设置在具有例如图 7 所示的出站和进站通路的信道配置中，可是该系统可能只包括一个模块。

在这个实例中，主控制器和模块与作为用来管理系统中组件的串行总线系统、称作“SMBus”的系统管理总线互连。但是，SMBus的使用不是本专利的发明原理所必需的，以及可采用组件之间的其它通信形式，其中包括存储信道通路本身。

5 根据本专利的发明原理、用于在示范系统中检测和标出故障位通道的方法的一个实施例如下所述。主控制器在出站通路的各位通道上发送测试模式。测试模式由各模块上的缓冲器接收并沿出站通路再驱动，直到它到达最外面的模块。最外面的模块则在进站通路的各位通道上发送测试模式。测试模式由各模块上的缓冲器接收并沿进站通路
10 再驱动，直到它到达主控制器。主控制器和模块上的缓冲器观察进站和/或出站通路的各位通道上的测试模式，以便检查正确的位通道操作。进站和出站通路上的位通道可同时被测试。

通过经由 SMBus 向主控制器发送结果和/或通过经由信道向主控制器传送结果帧，来报告故障位通道。这种结果帧可由最外面的模块
15 在进站通路上发起，以及其它模块(若有的话)可将它们的结果信息合并到进站通路中的数据中。如果来自各模块的结果在一个以上位通道上以冗余方式传送，则故障位通道不可能干扰报告结果。

一旦主控制器接收到结果，它可通过 SMBus、通过信道或者通过其它任何形式的通信向模块发出配置命令。配置命令向模块指示哪
20 些(若有的话)位通道是不良的且应当被标出。模块通过操作一个或多个故障在线恢复电路以绕过不良位通道(若有的话)重定向信号，并且重新配置任何内部功能性以适应位通道的丢失，来响应该配置命令。例如，如果一个位通道被指定用于对数据进行检错，则缓冲器或模块可禁用检错功能。

25 以上所述的故障在线恢复方法及设备的实施例仅作为示范，并且本专利的发明原理不限于这些具体实例。参照具有例如图 3 的实施例的单独的进站和出站通路的存储系统描述了根据本专利的故障在线恢复方法及设备的原理，但这些原理还可适用于采用单向链路的任何

存储体系结构，例如利用链路的环型布置、如 RamLink 的体系结构。

本专利的发明原理的一部分涉及置换状态模式。在其中存储器读和写数据在存储代理之间传递的、例如以上参照图 1 和图 3 所述的那些存储系统中，在存储代理之间发送状态信息、如空闲模式、告警模式及其它状态信息可能也是有用的。这可通过在连接存储代理的一个或多个相同链路上发送数据模式和状态模式来实现。根据本专利的发明原理，状态模式可随时间而置换。

例如，参照图 3，存储控制器 50 可向模块 52 中的一个或多个发送具有数据模式、如读命令的帧，模块 52 中的一个或多个可通过向控制器发回具有数据模式、如读数据的帧来响应。例如，如果模块不能够足够快地从存储装置 58 中检索读数据，则让模块中的一个或多个向存储控制器发回具有空闲模式的帧可能是有用的。可指定预定数据模式为空闲模式，使得如果存储控制器接收到空闲模式，它知道不是在接收读数据。但是，如果实际读数据模式正好匹配指定的空闲模式，则这可能产生问题。

根据本专利的发明原理，存储控制器和一个或多个模块均可以能够以可预测方式置换空闲模式，使得空闲模式随时间而改变。例如，每当空闲帧被发送和/或接收时，存储控制器和模块可按照预定顺序来改变空闲模式。根据本专利的发明原理的这种方法的一个实施例如图 18 所示。这样，如果存储控制器发送读命令帧(158)并接收具有当前空闲模式(162)的响应帧(160)，则它可重发相同的读命令(164)。如果第二响应帧(166)包含与第一个相同的模式(168)，则它将该模式解释为实际读数据(170)。但是，如果第二响应帧中的模式匹配置换空闲模式(168)，则存储控制器知道，第一响应帧为空闲帧(172)。

根据本专利的发明原理，状态模式中发送的状态信息可能是空闲模式、告警模式以及其它状态信息，例如来自模块的命令错误信息、来自模块的热过载信息以及表明模块检测到存储信道的出站通路上另一个模块的存在的信息。一些类型的状态模式可实现为互补模式。

例如,告警模式(可用于向代理通知错误状况)可实现为空闲模式的逻辑互补。这可通过例如允许存储代理对空闲和告警模式使用相同的模式发生器来简化实现。即使没有使用置换模式,互补状态模式的使用也可能是有益的。

5 根据本专利的发明原理的存储代理还可以能够在状态模式中故意产生错误、如循环冗余校验(CRC)错误。这种技术作为区分数据模式与状态模式的替代或补充方式可能是有用的。例如,在一些存储系统中,各帧与用于检验帧中数据的完整性的CRC代码一起被发送。

10 根据本专利的发明原理,存储代理可故意随包含状态模式的帧一起发送错误的CRC代码。接收代理则可把该帧解释为状态帧而不是数据帧。一些存储系统可利用具有额外位通道的一个或多个通路来传送CRC数据。如果这种系统能够以故障在线恢复模式工作,则一个或多个代理在没有以故障在线恢复模式工作时可以仅利用故意CRC错误。本文所使用的术语“CRC”不仅表示循环冗余校验,而且还表示用于检验帧或模式的完整性的其它任何类型的检错方案。

15 虽然根据本专利的发明原理的状态模式置换和处理的原理适用于任何类型的存储代理并且与本专利的其它发明原理无关,但将针对例如图7所示的实施例的存储缓冲器以及在例如图6所示的实施例的系统的上下文中描述一些附加方面。参照图6,如果存储缓冲器64
20 是存储信道上最外面的代理,则每当它没有发送主控制器从接到存储接口68的任何存储装置请求的数据时,它可以能够在进站链路56B上不断传送置换空闲状态帧。

25 图19说明根据本专利的发明原理的置换模式发生器的一个实施例。图19的实施例是具有多项式 $x^{12}+x^7+x^4+x^3+1$ 的12位线性反馈移位寄存器(LFSR)。初始状态可设置为000000000001,以及LFSR在模式重复之前循环通过 $2^{12}-1$ 种状态(4095帧)。LFSR的每个位可映射到数据通路上的链路中的位通道,以及每个位可用于在整个帧期间在相应位通道上出现的所有传递。例如,在具有在各链路中带12个位通

道的数据通路的系统中，来自 LFSR 的每级的输出可映射到位通道之一。通过利用来自延迟一帧的 LFSR 的最低有效位的值，可容纳附加通道、例如第 13 个位通道。

5 图 20 说明图 19 的置换模式发生器产生的第一状态模式的一个实例。在这个实例中，帧为 12 个传递长。图 21-22 分别说明第二、第三和第四状态模式。通过在整个帧期间使用各位通道上的相同值，电磁干扰(EMI 或噪声)可减小。

10 这里所述的 13 位通道 \times 12 位传递帧是作为实例，本专利的发明原理不限于这些细节，也不限于上述置换模式发生器的具体实施例。例如，根据本专利的发明原理的置换模式发生器不需要采用专用逻辑电路、如上述 LFSR 来实现。作为选择，它可采用可编程逻辑、或作为处理器中的算法或者可用来监视和/或实现存储接口中逻辑或利用置换状态模式的缓冲器或其它存储代理的其它功能性的其它可编程状态机来实现。

15 本专利的一些附加发明原理涉及利用一个以上位通道来检测存储链路上是否存在存储代理。例如，在图 7 所示的存储缓冲器的实施例中，缓冲器可以能够检测是否存在耦合到出站链路 54B 的另一个存储代理。这可通过利用链路中的单个位通道测试另一个存储代理的存在来实现。但是，如果链路中存在一个以上位通道，则根据本专利的发明原理，位通道中不止一个可用来检测另一个存储代理的存在。
20 这可防止不良位通道的存在干扰存在检测操作。

为了方便，涉及利用一个以上位通道来检测存储代理的存在的本专利的发明原理将单独和统一地称作冗余存在检测。冗余存在检测可适用于具有与多个位通道的链路接口的任何类型的存储代理。例如，
25 图 13 的实施例中所示的发送器 118 中的任何两个或两个以上可看作是链路接口，在本例中为发送链路接口。同样，图 13 所示的接收器 116 中的任何两个或两个以上可看作是链路接口，在本例中为接收链路接口。冗余存在检测可适用于这些链路接口中任一个以及图 17 的

实施例所示的链路接口 126 和 128 中任一个。

再回到图 7 的实施例作为实例，当存在检测事件、如重置出现时，存储缓冲器可将其进站发送链路 56B 上的三个位通道驱动到预定存在检测逻辑电平、例如一，从而向另一个缓冲器发信号通知其存在。同样在存在检测事件中，在信道上位于第一缓冲器的进站处的第二个这种存储缓冲器可将其进站接收链路 56A 上相应的三个位通道配置成检测第一缓冲器的存在。在这个实例中，第一存储缓冲器将称作外部代理，而第二缓冲器则称作内部代理。

将位通道配置成检测另一个代理的存在的技术的一个实例将让那个位通道的接收器尝试把偏置电流施加到位通道上，从而迫使该位通道与存在检测逻辑电平相反。如果另一个存储代理在存在检测事件期间耦合到该位通道，则它在那个位通道上的发送器将迫使位通道达到存在检测逻辑电平。

如果内部代理检测到三个位通道中的两个上的存在检测逻辑电平，则它知道外部代理存在，并且它可让其外部端口的全部或一部分被启用。(在这个实例中，外部端口包括用于出站链路 54B 的链路接口和用于进站链路 56A 的链路接口。)如果内部代理无法检测到三个位通道中的至少两个上的存在检测逻辑电平，则它可判定外部代理不存在，并且它可禁用其外部端口的全部或一部分。内部代理可以能够响应状态检查命令而向另一个代理、例如向存储控制器报告外部代理是否存在。

利用冗余存在检测的代理还可以能够向另一个代理发信号通知存在检测事件。例如，如果重置事件在出站通路上通过重置命令传递到图 7 的缓冲器，则这个命令可由再驱动电路 60 转发给外部代理(若有的话)。这可将两个代理设置为存在检测模式。

根据本专利的发明原理的冗余存在检测不限于上述具体实施例。例如，只有两个位通道可用于存在检测而不是以上实例中的三个，在这种情况下，内部代理只需要检测单个位通道上的存在检测逻辑

辑电平以推断外部代理存在。同样，冗余存在检测可适用于利用其它各种类型的存储体系结构的系统和组件，例如利用链路的环型布置、如 RamLink 的体系结构。

5 根据本专利的一些附加发明原理涉及组件从存储信道中热插和/或热拔—即在存储信道正工作时增加和/或移去组件。图 24 说明根据本专利的发明原理的存储代理 134 的一个实施例。图 24 的实施例可以是存储模块、存储缓冲器、存储控制器等。代理包括第一端口 136 和第二端口 138。只是为了说明，如果假定代理为存储模块、例如图 6 的实施例中的模块 52 之一，则第一端口可指定为内部端口，因为它可设置成与存储信道上更接近存储控制器的其它代理进行通信。同
10 样，第二端口可指定为外部端口，因为它可设置成与存储信道上远离存储控制器的代理进行通信。这些指定只是用于说明，并且这些发明原理不限于存储代理的这些细节，也不限于图 6 所示的存储信道的具体情况。这些原理还可适用于其它存储信道体系结构，如图 1 所示的
15 RamLink 体系结构。

根据本专利的发明原理的存储代理的每个端口具有一个或多个链路接口。在图 24 的实施例中，每个端口具有接收链路接口和发送链路接口。内部端口 136 具有可能是属于再驱动电路 60 的一部分的一个或多个接收器的接收链路接口 140 以及可能是属于另一个再驱动电路 62 的一部分的一个或多个发送器的发送链路接口 142。外部
20 端口分别具有同样分别是再驱动电路 62、60 的组成部分的接收和发送链路接口 144、146。链路接口 140、146 可分别耦合到出站链路 54A、54B，以及链路接口 142、144 可分别耦合到进站链路 56B、56A。链路接口的每个可具有一个或多个位通道，以及位通道和接口可采用
25 这个术语的任何组件来表示。例如，接口 142 中的位通道可表示为进站发送或进站 Tx 位通道。接口 144 中的位通道可表示为进站接收或进站 Rx 位通道。

图 24 的实施例仅作为示范，以及存储代理和端口可按照不同方

式来实施。例如，链路接口不一定是再驱动电路的组成部分。这种情况在图 17 所示的存储控制器的实施例中说明，其中，端口可包括不属于再驱动电路的组成部分的链路接口 126 和 128。链路接口可以仅包括一个或者任何数量的位通道，以及端口可以仅具有一个接收链路接口或发送接口。

5 根据本专利的发明原理的存储代理可以能够检测其端口之一上的另一个存储代理的存在，并且它可以能够根据另一个存储代理是否存在来采取各种动作。例如，图 24 的存储代理可以能够当另一个存储代理在其外部端口上不存在时禁用该端口的全部或一部分。它可以能够通过其内部端口向另一个代理、例如向存储控制器报告外部代理是否存在。图 24 的存储代理可以能够执行存在检测操作，它可包括向外部端口上潜在的外部代理发信号通知存在检测事件。它还可以能够执行快速重置操作。

10 将在存储系统的一个示例实施例的上下文中描述根据本专利申请、可帮助热添加/移去的一些附加发明原理。将在例如图 6 的实施例的存储系统的上下文中，参照图 24 的存储代理来描述该示例实施例。在这个示例实施例中，假定图 24 的存储代理用来实现作为具有存储装置的模块的一部分的图 6 的缓冲器的一个或多个。但是，所有这些细节只是用于说明，以及发明原理不限于这些细节。

20 在示例系统中，存储代理可以能够执行快速重置操作、完全重置操作和/或各种轮询或存在检测操作。在示例系统中，最小数量的时钟转变可能是使各位通道上的导出时钟保持锁定到数据流所必需的。因此，存储控制器(或主控制器)可通过在预定时间周期在出站通路中的位通道中一个或多个上发送一或零的连续流来发起重置操作。由于数据由通路上的各缓冲器再驱动，因此所有缓冲器接收重置命令或事件。在示例系统中，三个最低有效位(LSB)通道可用来发送重置操作的信号。接收代理可通过在三个 LSB 中任何两个上检测零或一的流来检测重置事件。这可保证一个故障位通道的存在不干扰重

25

置操作,但甚至不要求多于一个位通道的发明原理不限于这样一种实现。

在示例系统中,例如在主控制器通过外部条件保持为重置时,主控制器可发送零的连续流,使信道上的所有代理(在这个实例中为具有缓冲器的模块)无限地保持第一重置状态。主控制器则可在第一时间量、例如两个帧周期发送一的流,然后重新发送零,以便发信号通知其它代理执行快速重置操作。或者,主控制器可在第二时间量、例如不止两个帧周期发送一的流,以便发信号通知其它缓冲器执行完全重置操作。完全重置可包括各种内部校准操作,例如链路上的阻抗匹
5 配、任何接收器或驱动电路中的电流源校准、接收器偏移抵消等。在执行校准操作之后,主控制器则可发信号通知缓冲器转变到快速重置操作。

快速重置操作可绕过完全重置过程中执行的某些操作、如校准操作。快速重置操作可从存在检测操作开始。在存在检测操作过程中,信道上的各缓冲器可在三个 LSB 进站 Rx 位通道上施加电流,从而迫使这些位在没有连接到外部代理时为零。另外在存在检测操作过程中,各缓冲器还可把三个 LSB 进站 Tx 位通道驱动到一。每个缓冲器则可检查其三个 LSB 进站 Rx 位通道,以及如果它检测到三个通道其中两个上的一,则可让其外部端口保持启用,并相应地更新状态寄存器。如果缓冲器没有检测到两个一,则它可假定没有外部代理,禁用
15 其外部端口的全部或一部分,将其本身配置成执行信道上最外面的代理的功能,和/或相应地更新状态寄存器。主控制器可按照类似的存在检测操作来确定任何代理是否处于信道上。缓冲器可响应来自主控制器的状态请求而在状态帧中把状态信息转发给主控制器。

在存在检测操作之后,示例系统中的缓冲器可在快速重置过程中经过其它各种操作,例如训练缓冲器上的本地时钟以锁定到数据流的时钟训练状态、对准信道上发送的帧的帧训练状态、检查所有位通道的操作并在缓冲器具有这种功能时将其设置到故障在线恢复模式的
25

位通道测试等。另外，一旦主控制器知道其它代理连接到信道的方式，它可调整帧大小、定时等，以接纳所有代理。

在示例系统中，存储代理还可以或者备选地能够执行各种轮询操作，以便检测信道上新增代理的存在。例如，各缓冲器在作为确定是否已经向信道添加新代理的最外面缓冲器时可以能够对其外部端口
5 执行轮询操作。图 25 说明根据本专利的发明原理的这种轮询操作的一个实施例。

在 148，代理可禁用其外部端口的全部或一部分。如果代理是缓冲器或模块，则可等待来自主控制器的轮询命令以便转变到 150 的热
10 重置操作。如果代理为主控制器，则它可禁用其外部端口的全部或一部分，并等待来自系统环境的唤醒命令。在接收到唤醒命令时，它可转到启用其外部端口的全部或一部分，并转变到重置状态。

在 150，代理可启用其外部端口，并把零驱动到三个 LSB 出站 Tx 位通道，以便向其外部端口上潜在的新代理发送重置。代理则可
15 转变到 152 的热校准操作。

在 152，代理可把一驱动到三个 LSB 出站 Tx 位通道，以便迫使潜在的新代理通过包括校准操作的完全重置，因为新检测的代理可能需要被校准。然后，代理可转变到 154 的热检测操作。

在 154，代理可将零驱动到三个 LSB 出站 Tx 位通道，并且在三个
20 个 LSB 进站 Rx 位通道上施加偏置电流，从而迫使这些位在没有连接到外部代理时为零。代理则可检查三个 LSB 进站 Rx 位通道，以及如果它检测到至少两个一，则它可在 155 判定外部代理存在，并转变到 156 的热代理存在操作。否则，代理可在 155 判定外部代理不存在，并重新转变到 148 的睡眠操作。

在 156，代理可更新状态寄存器以表明它已经检测到外部代理，
25 然后例如响应状态请求而把这个信息转发给主控制器，或者采取一些其它动作把信息转发给主控制器或其它代理。代理也可等待接收信道重置。

主控制器可通过定期状态请求或其它技术变为知道新检测的代理，并发起快速重置，以使用信道上的新代理对整个信道重新初始化并适应信道定时。

5 下面是根据本专利的发明原理的热增加/移去程序的一些附加实施例。还参照图 6 所示的存储系统的实施例，在采用图 6 的存储系统的更大系统、如具有用户界面和系统固件的服务器的上下文中描述这些附加实施例。但是，这些附加实施例所述的发明原理不限于本文所述的具体细节。

10 根据本专利的发明原理的热增加程序可在用户把新代理附加到存储信道上、例如最外面代理的外部端口上时开始。用户可通知系统固件已经附加代理。固件则可使电力施加到附加代理上，并通过唤醒命令通知主控制器已经附加代理。主控制器则可向先前的最外面代理发送轮询命令，先前的最外面代理则可循环进行例如以上参照图 25 所述的轮询操作。在轮询操作之后，先前的最外面代理可报告新的外部代理的存在。主控制器则可检测新代理的存在，并发出快速重置命令以便使新代理投入运行，并对整个信道重新定时。在新代理可工作之后，主控制器可中断系统固件以报告新代理是可工作的。或者，主控制器可等待系统固件查询主控制器以确定新代理是否可工作。然后，系统固件可把主控制器配置成适应新代理提供的任何新硬件、如新的存储装置，它在代理为存储模块或缓冲器时可能存在。

25 根据本专利的发明原理的热移去程序可在用户通知系统要移去存储信道上的特定代理时开始。系统可从系统图中消除相应的主控制器地址范围。如果系统使用镜像，则系统可能把主控制器地址范围重新映射到代理镜像。然后，如果还没有镜像，则系统可从主控制器地址范围中把数据复制或移动到其它位置。系统则可进行轮询，直到完成所有未决事务。然后，系统可使主控制器向要移走的代理内部的代理发送命令，使这个代理认为它是信道上最外面的代理，从而使它禁用其外部端口，并在后续快速重置过程中承担最外面代理的功能。(完

全重置将忽略这个命令。)然后,系统可发起快速重置以关闭所选代理以及附加到所选代理的组件的任何信道接口。然后,系统可断开所选代理的电力,并通知用户已经移去该代理。

5 根据本专利的发明原理的热替换程序可在上述热移去程序完成时开始。用户可添加新代理来取代移去的代理,然后通知系统固件已经添加新代理。正运行的系统则可为新替换组件预备主控制器,并向新组件供电。然后,系统固件可使主控制器向先前的最外面代理发送命令,让它知道它应当不再认为它是最外面的代理。这可使先前的最外面代理响应下一个重置而启用其外部端口,并等待轮询命令。固件
10 则可指示主控制器向先前的最外面代理发送轮询命令,先前的最外面代理则可执行例如以上参照图 25 所述的轮询操作,从而初始化新的代理。然后,先前的最外面代理可报告新的外部代理的存在。主控制器则可检测新代理的存在,并发出快速重置命令以使新代理投入运行,并对整个信道重新定时。在新代理可工作之后,主控制器可中断
15 系统固件以报告新代理是可工作的。或者,主控制器可等待系统固件查询主控制器以确定新代理是否可工作。

可在设置和细节方面对本文所述的实施例进行修改,而没有背离发明原理。因此,这类变更和修改被视作落入以下权利要求的范围之内。

20

图 1
(现有技术)

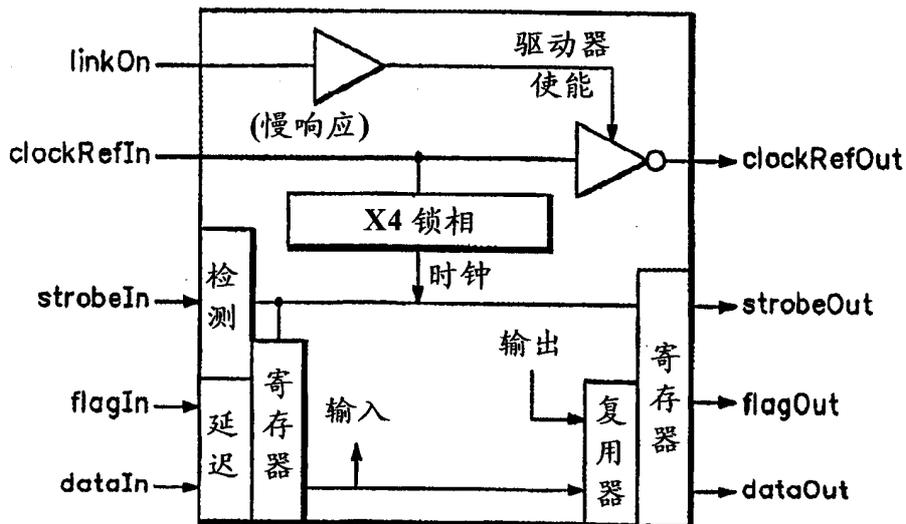
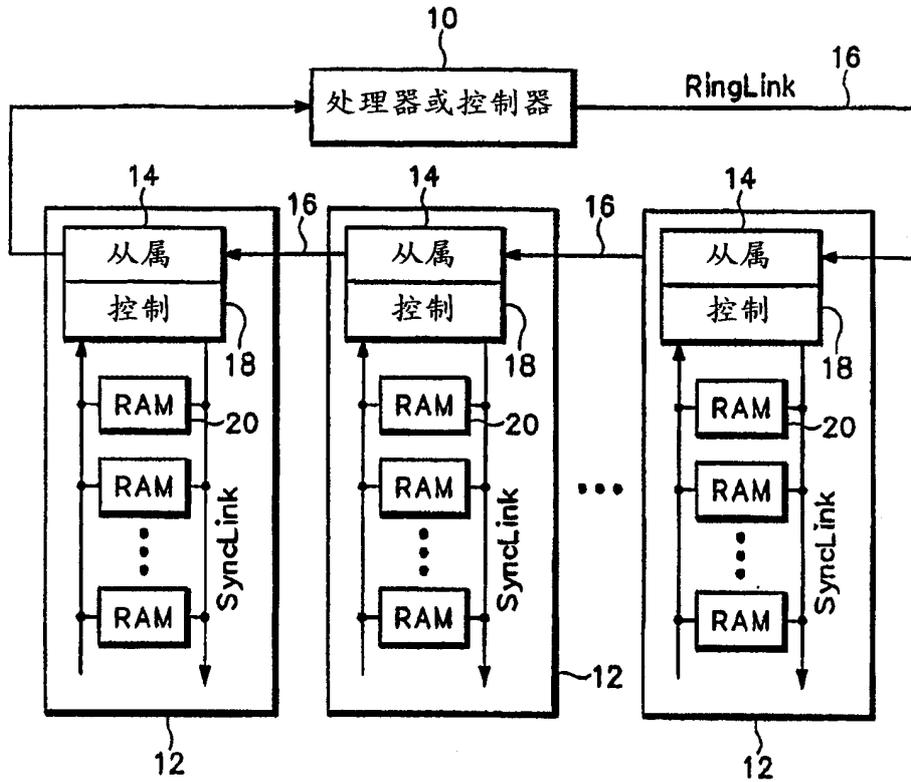


图 2
(现有技术)

图 3

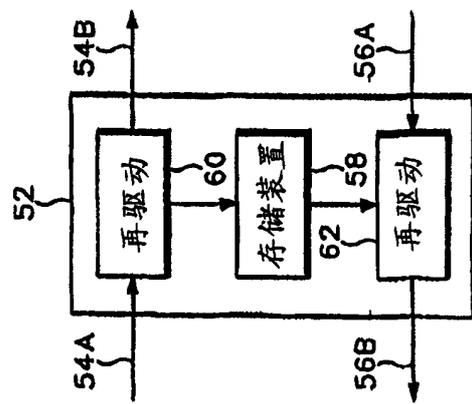
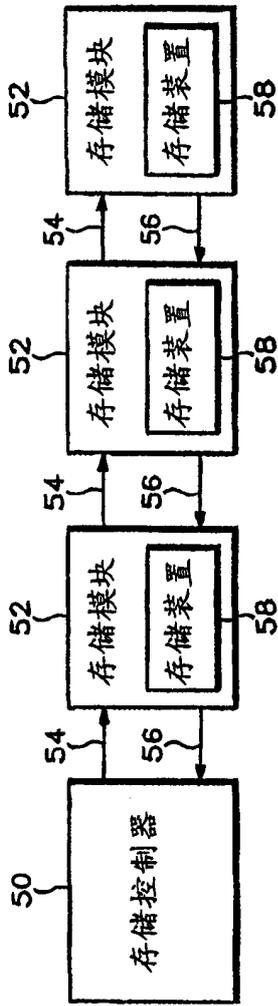


图 4

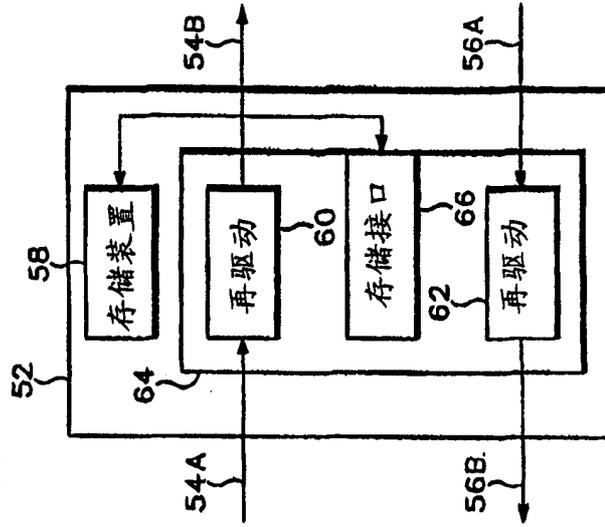


图 5

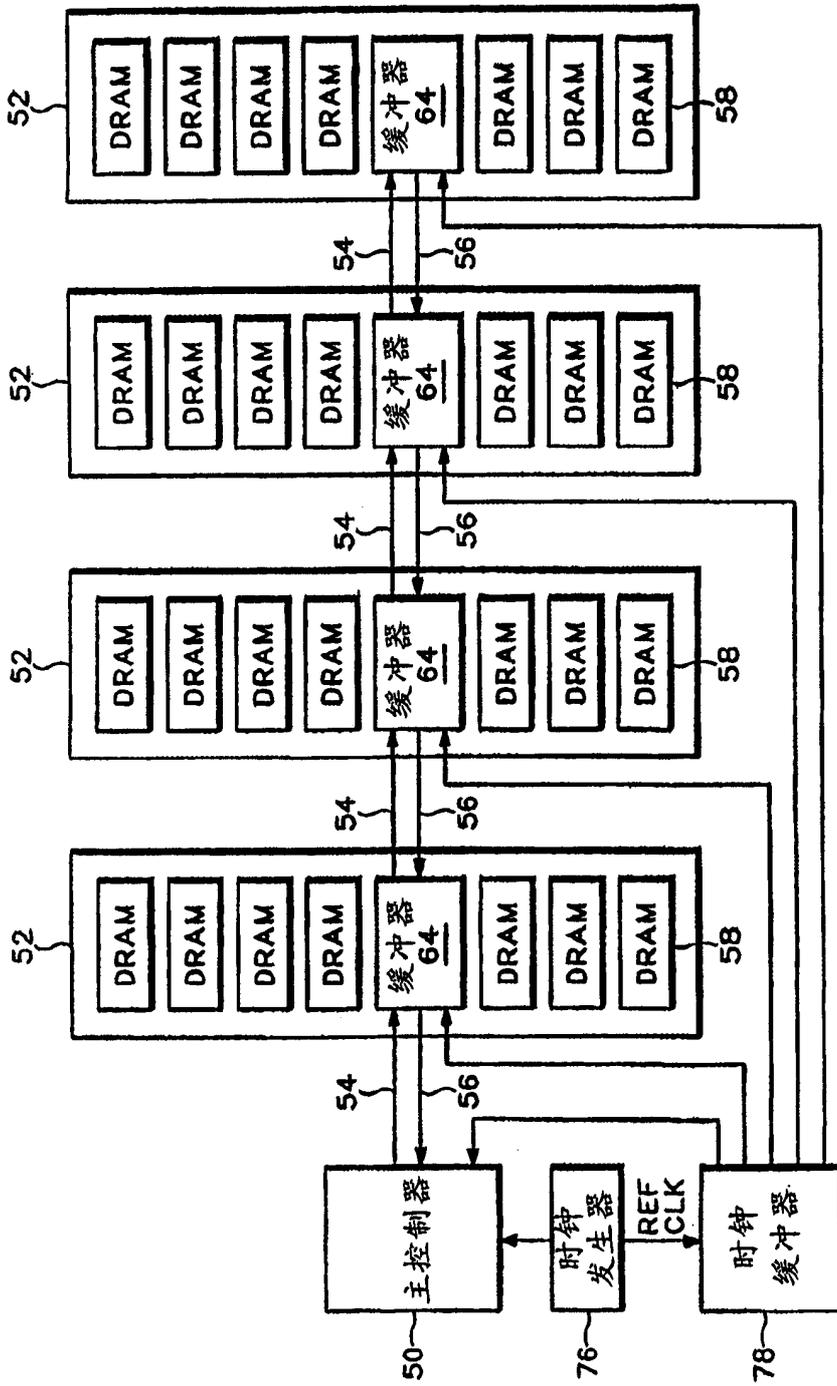


图6

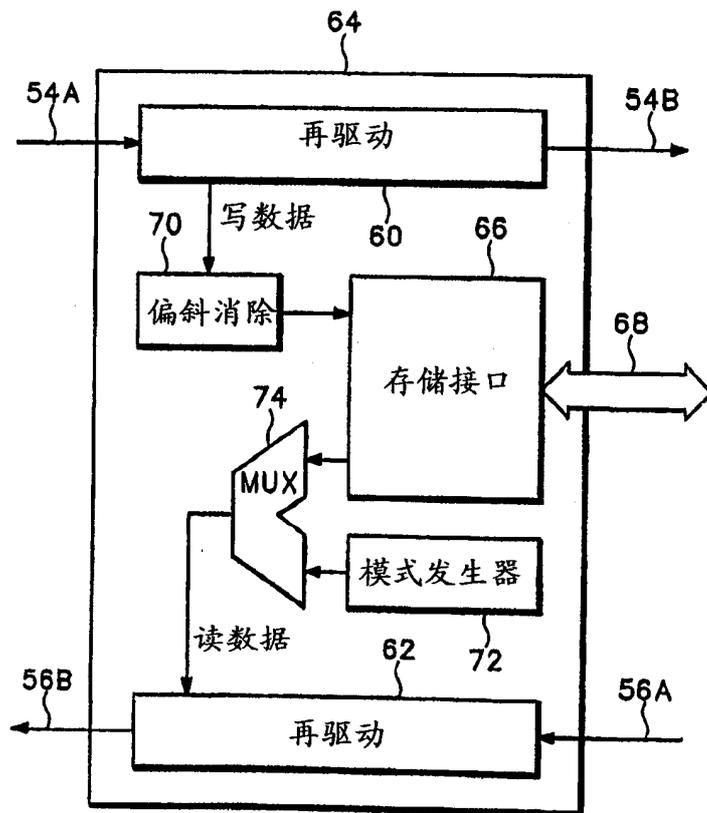


图 7

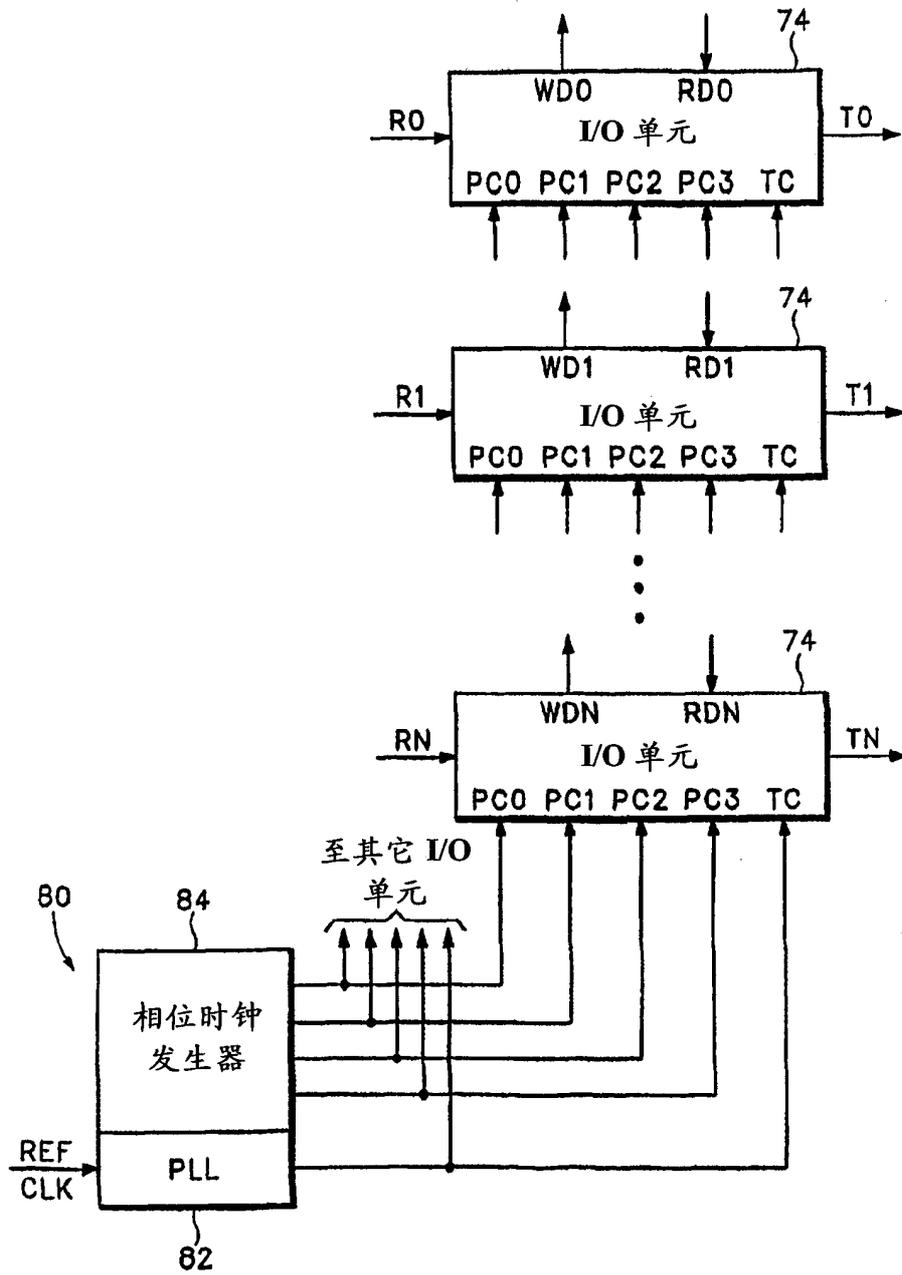


图 8

图 9

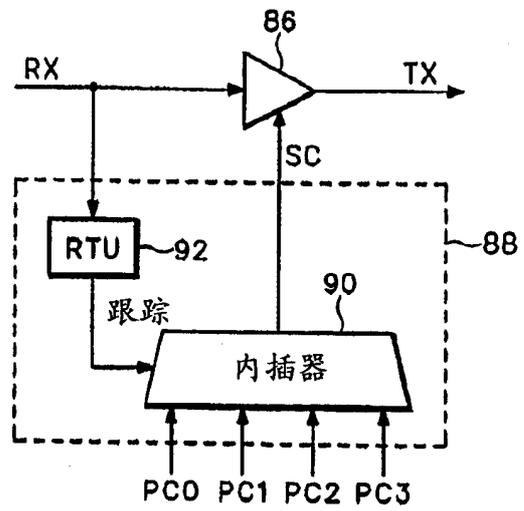
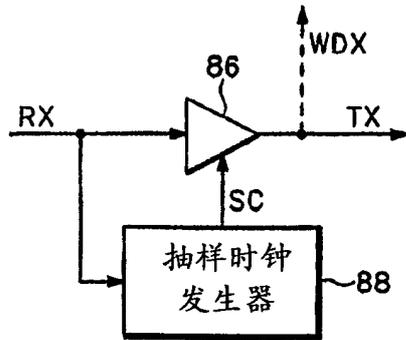


图 10

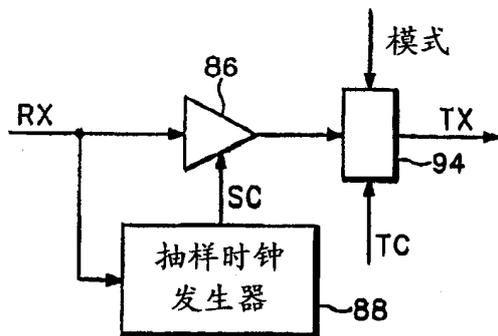


图 11

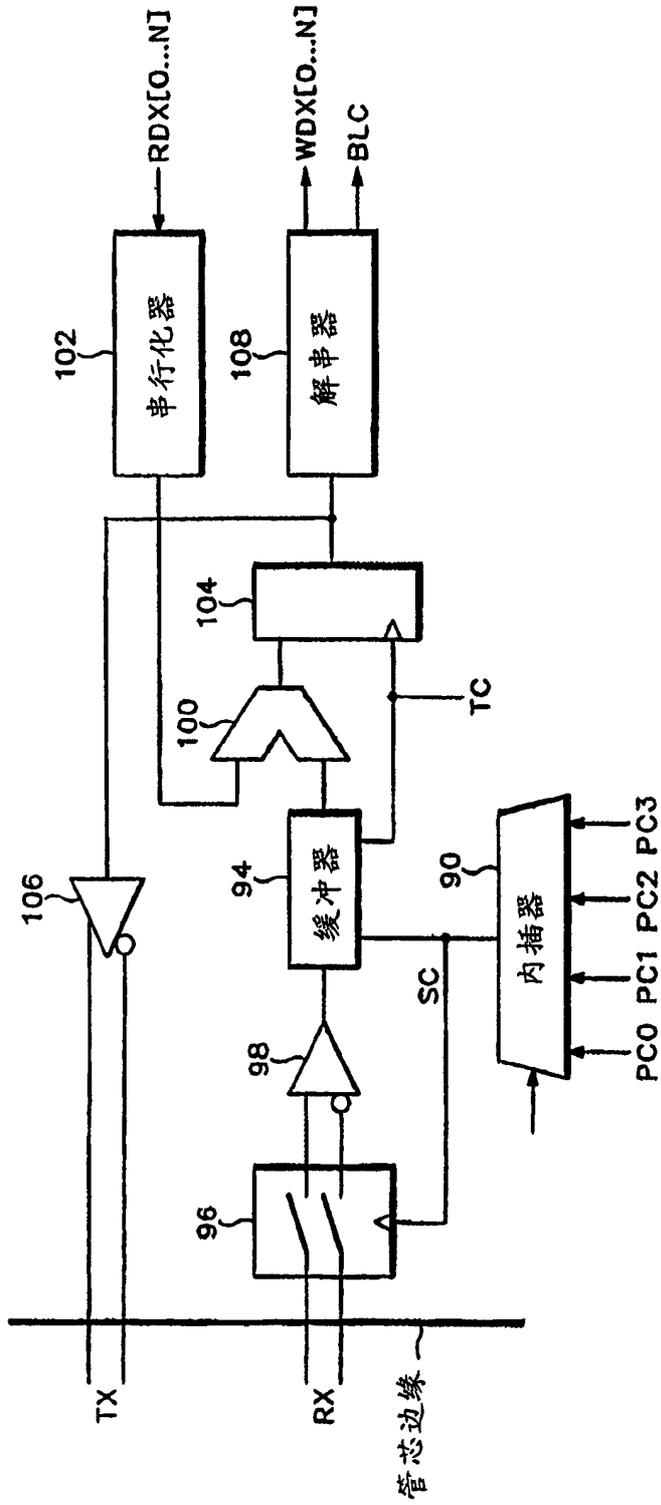


图 12

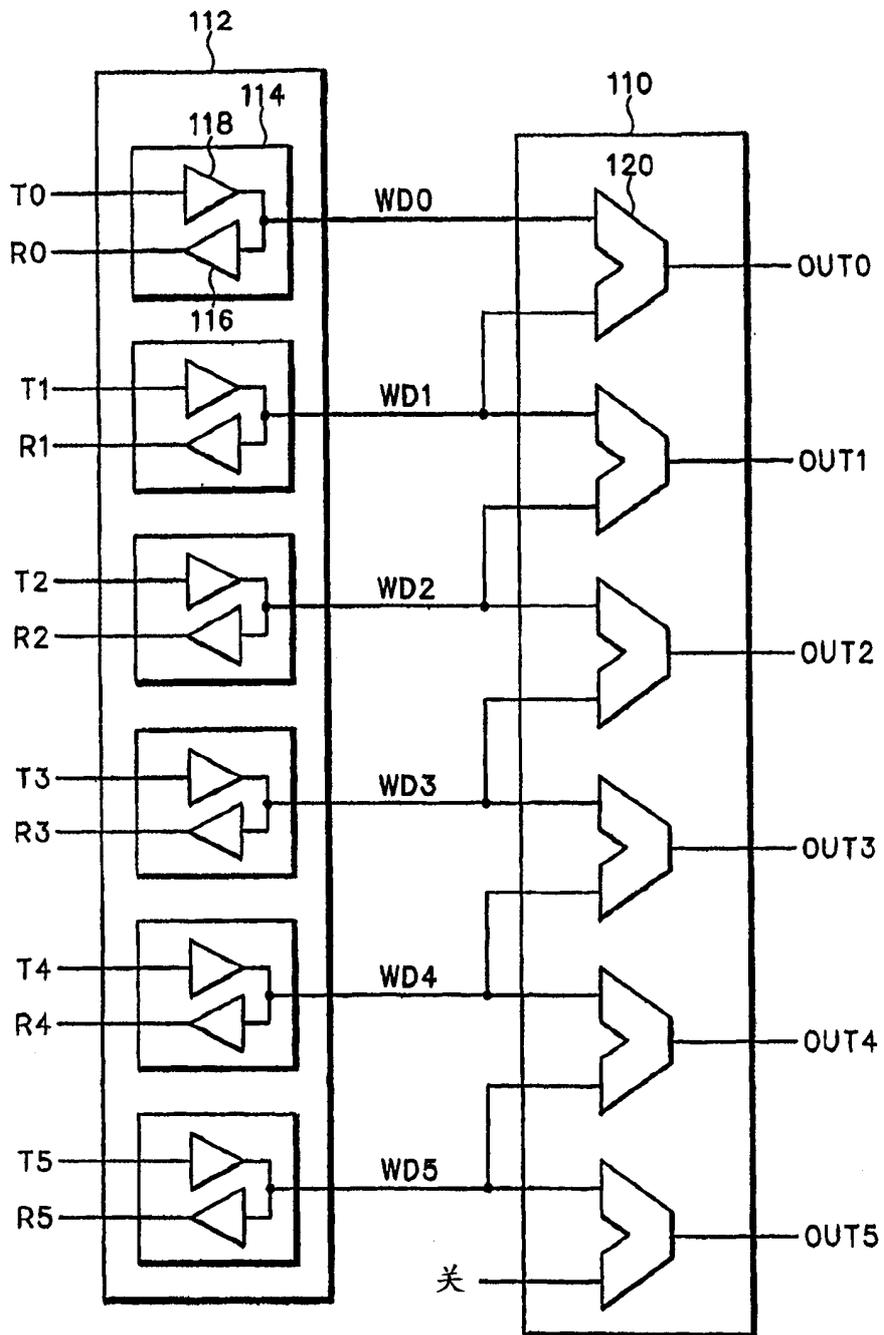


图 13

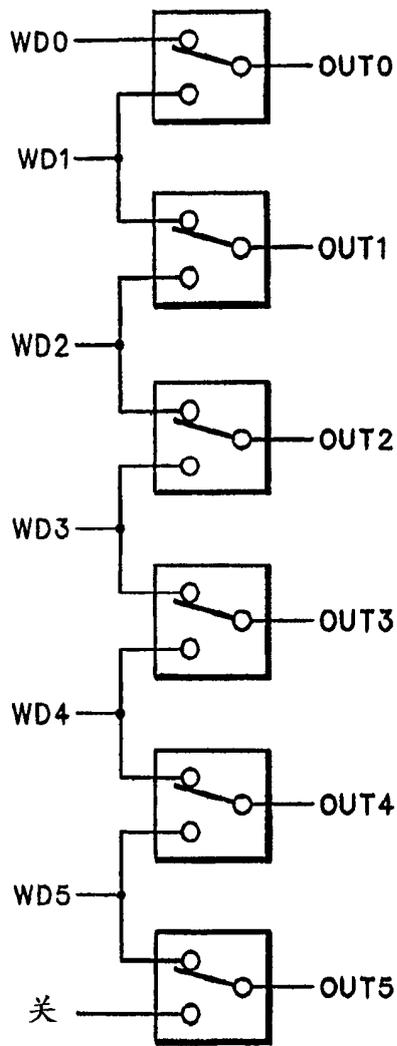


图 14

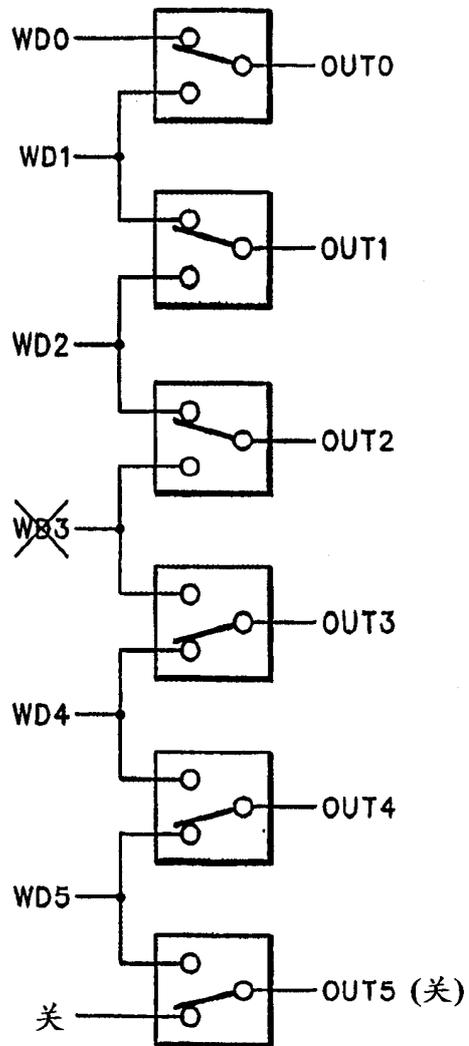


图 15

图 16

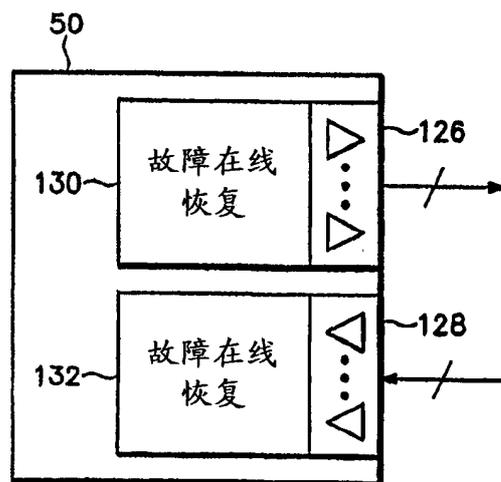
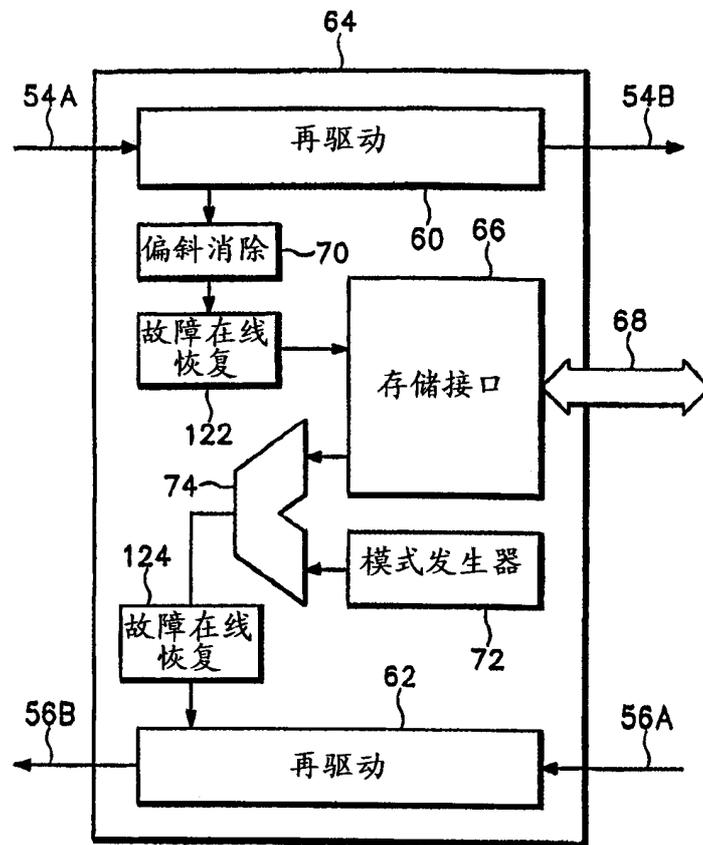


图 17

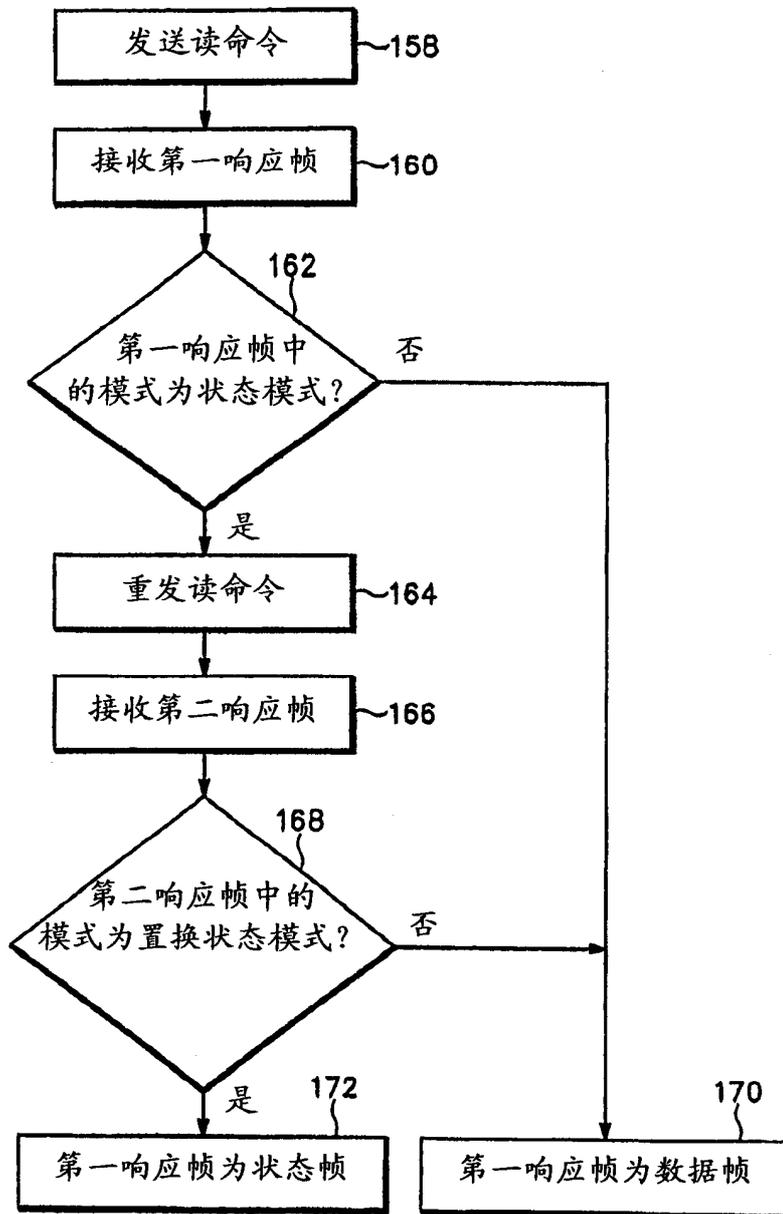
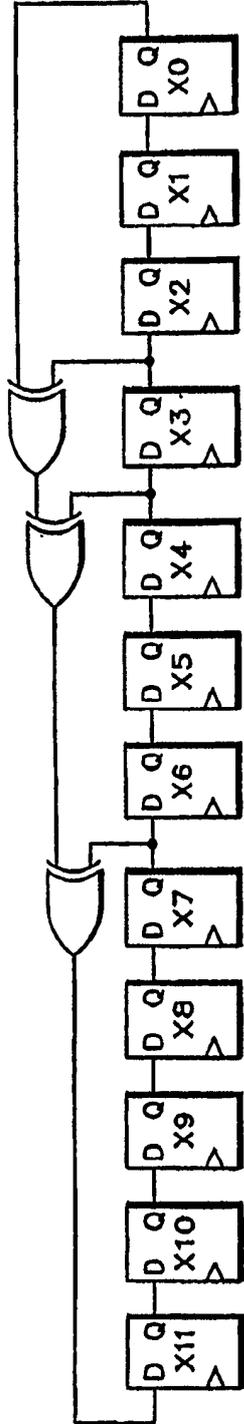


图 18

图 19



位	12	11	10	9	8	7	6	5	4	3	2	1	0
传递	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	1

图 20

图 21

传递	位	12	11	10	9	8	7	6	5	4	3	2	1	0
0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1	0	0	0	0	0	0	0	0	0	0	0	0
5	1	1	0	0	0	0	0	0	0	0	0	0	0	0
6	1	1	0	0	0	0	0	0	0	0	0	0	0	0
7	1	1	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	0	0	0	0	0	0	0	0	0	0	0	0
9	1	1	0	0	0	0	0	0	0	0	0	0	0	0
10	1	1	0	0	0	0	0	0	0	0	0	0	0	0
11	1	1	0	0	0	0	0	0	0	0	0	0	0	0

传递	位	12	11	10	9	8	7	6	5	4	3	2	1	0
0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
2	0	0	1	0	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0
4	0	0	1	0	0	0	0	0	0	0	0	0	0	0
5	0	0	1	0	0	0	0	0	0	0	0	0	0	0
6	0	0	1	0	0	0	0	0	0	0	0	0	0	0
7	0	0	1	0	0	0	0	0	0	0	0	0	0	0
8	0	0	1	0	0	0	0	0	0	0	0	0	0	0
9	0	0	1	0	0	0	0	0	0	0	0	0	0	0
10	0	0	1	0	0	0	0	0	0	0	0	0	0	0
11	0	0	1	0	0	0	0	0	0	0	0	0	0	0

图 22

图 23

传递 \ 位	12	11	10	9	8	7	6	5	4	3	2	1	0
0	0	0	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0
2	0	0	0	1	0	0	0	0	0	0	0	0	0
3	0	0	0	1	0	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0	0	0
5	0	0	0	1	0	0	0	0	0	0	0	0	0
6	0	0	0	1	0	0	0	0	0	0	0	0	0
7	0	0	0	1	0	0	0	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0	0	0	0	0	0
9	0	0	0	1	0	0	0	0	0	0	0	0	0
10	0	0	0	1	0	0	0	0	0	0	0	0	0
11	0	0	0	1	0	0	0	0	0	0	0	0	0

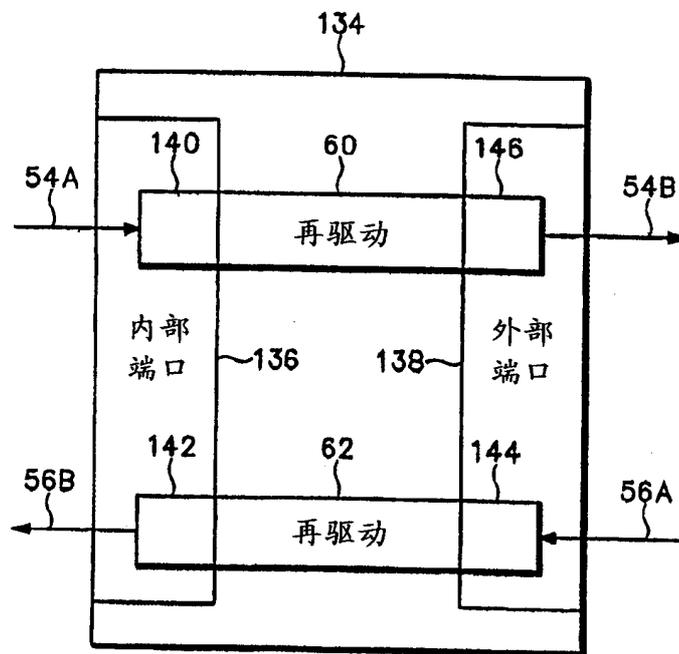


图 24

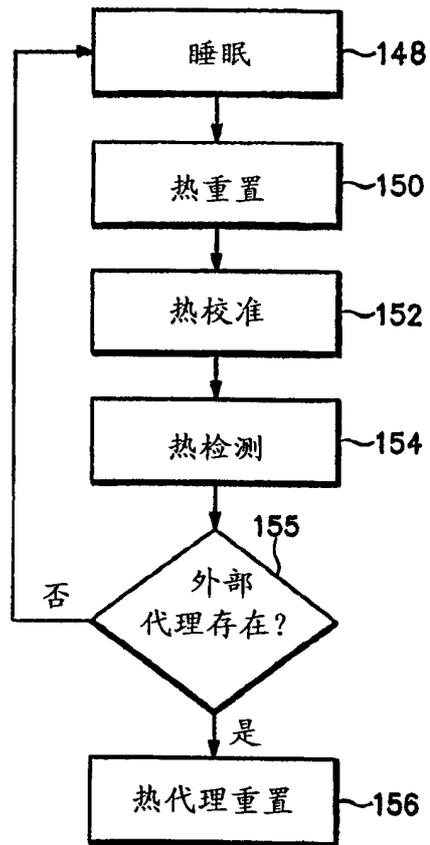


图 25