



[12] 发明专利说明书

专利号 ZL 200610135528.5

[45] 授权公告日 2009 年 6 月 24 日

[11] 授权公告号 CN 100504749C

[22] 申请日 2006.10.18

[21] 申请号 200610135528.5

[30] 优先权

[32] 2005.10.20 [33] US [31] 11/254,635

[73] 专利权人 戴尔产品有限公司

地址 美国得克萨斯州

[72] 发明人 K·马可斯 J·切里安

[56] 参考文献

US6343324B1 2002.1.29

WO2004095255A2 2004.11.4

CN1205547C 2005.6.8

审查员 王燕_1

[74] 专利代理机构 北京纪凯知识产权代理有限公司

代理人 程伟 王锦阳

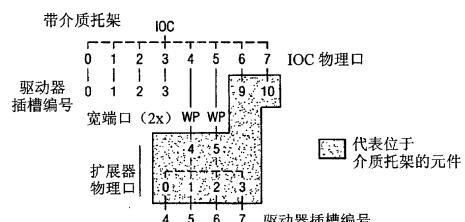
权利要求书 2 页 说明书 8 页 附图 4 页

[54] 发明名称

将串行附属 SCSI 控制器端口映射到磁盘驱动器的方法

[57] 摘要

一种用于数据存储系统的磁盘驱动器到服务器连接插槽的映射的方法。当使用 SAS 扩展器来添加额外的磁盘驱动器，并且维持相同的驱动器编号计划，就好像不存在扩展器的时候，可以使用该方法。该方法使用 SAS 连接的标识地址帧来确定一个设备是否连接到控制器端口的每个物理口，以及该设备是扩展器或是终端设备(磁盘驱动器)。



1. 一种在数据存储系统中、将串行附属 SCSI 控制器的端口映射到磁盘驱动器的方法，所述数据存储系统包括至少一个用于连接到一个或多个磁盘驱动器的扩展器，所述控制器包括用于 x 个连接的端口，即物理口 0— x ，该方法包括：

 在所述控制器的每个物理口上，接收串行附属 SCSI 标识地址帧；

 从所述控制器包含物理口 0 的端口开始，从所述标识地址帧读取附加设备类型；

 如果所述标识地址帧指出与所述端口关联的设备是一个终端设备，将设备 ID 0 指定给该设备；

 如果所述标识地址帧指出与所述端口关联的设备是一个扩展器，将设备 ID 0— x 指定给具有到所述扩展器的物理口的每个下一驱动器 1— x ；

 如果没有接收到标识地址帧，指出没有设备与所述端口关联，为物理口 0 保留设备 ID；以及

 为所述控制器的物理口 1— x 重复上述步骤，按照相关联的磁盘驱动器的发现次序指定连续的设备 ID。

2. 如权利要求 1 所述的方法，进一步包括基于所述设备 ID，控制与所述磁盘驱动器相关的状态 LED。

3. 如权利要求 2 所述的方法，其中 RAID 类型的驱动器控制器从所述控制器接收设备 ID 并且控制所述状态 LED。

4. 如权利要求 1 所述的方法，其中所述系统里任意设备的最低编号的驱动器插槽，连接到用于所述设备的最低编号的物理口。

5. 如权利要求 1 所述的方法，进一步包括将磁盘驱动器插入空余的驱动器插槽，以及将一个保留的设备 ID 指定给所述磁盘驱动器。

6. 如权利要求 1 所述的方法，其中所述控制器从所述标识地址帧的设备类型域接收数据。

7. 如权利要求 1 所述的方法，其中所述控制器接收标识地址帧数据，该数据作为对用于扩展器所包含物理口的串行管理协议指令的响应。

8. 如权利要求 1 所述的方法，其中所述串行管理协议指令是发现 DISCOVER 指令。

将串行附属 SCSI 控制器端口映射到磁盘驱动器的方法

技术领域

本发明涉及数据存储系统，具体地说，涉及使用 SAS 扩展器（expander）的数据存储系统。

背景技术

许多现今的计算机设计有“底板”（backplane），这是包含了其他电路板可以插入的插座的电路板。例如，在个人计算机中，这个底板就是包含了用于各扩展卡的插座的大电路板。

服务器类型的计算机通常配备有“硬驱动器底板”。多个“插槽”的每个都有一个用于硬驱动器数据存储设备的连接器。

计算机系统的底板只接受固定数量的设备。一种用于将设备连接到底板的流行接口，称为 SAS（串行附属 SCSI）。SAS 系统是点对点配置，该配置可以使用扩展器，作为发起者（initiator）设备（如主机）与目标设备（如外设和典型情况下的存储设备）之间的中继设备。该扩展器允许系统内一个或多个发起者可以有到一个或多个目标的连接。

与现今的其他外设接口一样，SAS 接口设计为允许“热插拔”。这是指当计算机还在运行的时候，向计算机添加或从计算机删除设备的能力，以便操作系统自动识别这个变化。

发明内容

本发明提供一种在数据存储系统中、将串行附属 SCSI 控制器的端口映射到磁盘驱动器的方法，所述数据存储系统包括至少一个用于连接到一个或多个磁盘驱动器的扩展器，所述控制器包括用于 x 个连接的端口，即物理口 0— x ，该方法包括：在所述控制器的每个物理口上，接收串行附属 SCSI 标识地址帧；从所述控制器包含物理口 0 的端口开始，从所述标识地址帧读取附加设备类型；如果所述标识地址帧指出

与所述端口关联的设备是一个终端设备，将设备 ID 0 指定给该设备；如果所述标识地址帧指出与所述端口关联的设备是一个扩展器，将设备 ID 0—x 指定给具有到所述扩展器的物理口的每个下一驱动器 1—x；如果没有接收到标识地址帧，指出没有设备与所述端口关联，为物理口 0 保留设备 ID；以及为所述控制器的物理口 1—x 重复上述步骤，按照相关联的磁盘驱动器的发现次序指定连续的设备 ID。

依照本发明的含义，这里说明的系统和方法用于将数据存储系统的磁盘驱动器映射到服务器连接插槽的方法。当有一个或多个 SAS 扩展器被用来添加额外的磁盘驱动器的时候，可以使用该方法，而且，该方法可以维持相同的驱动器编号计划，就好像这些扩展器不存在一样。该方法使用 SAS 链接的标识地址帧数据，以确定一个设备是否连接到控制器端口的每个物理口上，而且确定这个设备是扩展器还是终端设备（磁盘驱动器）。

该方法分配一个设备 ID（Device ID）给系统中的每个磁盘驱动器。这个设备 ID 可以用于磁盘管理功能，如控制位于硬磁盘底板上的多个 LED。

该方法是“持续的（persistent）”，其含义在于，不论该服务器的数据存储子系统是否有扩展器，驱动器插槽的这些设备 ID 都是同样的。

附图说明

参考以下的附图说明并将其与附图共同使用，对本发明及其优点可以获得更完整的理解，附图中类似的附图标记指示了类似的特征，在这里：

图 1 说明配置了包括一个 6 端口 SAS 扩展器的 SAS 数据存储的服务器系统 100 的相关部分。

图 2 说明这些磁盘驱动器如何被映射到如图 1 所示、但没有扩展器的系统的驱动器插槽上。

图 3 说明这些磁盘驱动器如何被映射到图 1 所示系统的驱动器插槽上。

图 4 说明配置了包括一个 12 端口 SAS 扩展器的 SAS 数据存储的

服务器系统 400 的相关部分。

图 5 说明这些磁盘驱动器如何被映射到如图 4 所示、但没有扩展器的系统的驱动器插槽上。

图 6 说明这些磁盘驱动器如何被映射到图 4 所示系统的驱动器插槽上。

具体实施方式

图 1 说明配置了 SAS 数据存储的服务器系统 100 的相关部分。出于说明的目的，系统 100 可以是任何配置了 SAS 数据存储的“信息处理系统”。一个“信息处理系统”可以包括一种手段或多种手段的集合，这些手段均具有可操作性以计算、分类、处理、传输、接收、重新获得、产生、交换、保存、显示、展示、检测、记录、复制、操作或使用用于商业、科学、控制或其他目的的任意形式的信息、情报或数据。

例如，一个信息处理系统可以是一台个人电脑，一个网络存储设备，或任意其他合适的设备，也可以在尺寸、形状、性能、功能和价格上各有不同。该信息处理系统可以包括随机存取存储器 (RAM)，一个或多个如中央处理器 (CPU)、硬件或软件控制逻辑等的处理方式，ROM，和/或其他类型的非易失性存储器。该信息处理系统的其他组件包括一个或多个硬磁盘，一个或多个用于与外部设备通信的网络接口，以及诸如键盘、鼠标和视频显示器等的各类输入输出 (I/O) 设备。该信息处理系统还可以包括一条或多条总线，均具有操作性以在各种硬件组件之间传送消息。

图 1 中，除了该数据存储系统之外，只明确地显示了系统 100 最基本的元件。该系统有至少包含了一个激活存储器的处理系统 109，处理器，以及用于内部和 I/O 通信的总线。依照 SAS 的术语集，处理系统 109 这里被称为“发起者”，通常是指数据存储服务器。

在图 1 的例子中，系统 100 有 10 个磁盘驱动器 101。驱动器 101 通过存储控制器 105 与系统 100 的其他部分进行数据通信。在本说明的例子中，控制器 105 是一个 RAID 控制器，它只允许服务器 100 仅仅能见到直接处理这些驱动器的控制器 101。RAID（独立磁盘冗余阵列）包括两个或更多并行工作的磁盘。在用户看来，这些磁盘只是一个驱动器，而且，提供了改进的性能或安全性（或两者兼而有之）。执行 RAID 功能和控制这些硬磁盘的软件，位于控制器 105 上。

控制器 105 包含 SAS I/O 控制器 (IOC) 106，该控制器遵照 SAS 的标准控制、控制去往和来自这些磁盘驱动器的输入和输出。控制器 105 通过一条如 PCI 类型总线的总线 104，与系统 100 的剩余部分通信。

8 个磁盘驱动器 101 与底板 103 相连。底板 103 有与之相关的底板控制器 103a。

系统 100 还有至少一个介质托架 (media bay) 底板 107。介质托架底板 107 被设计为允许安装额外的外设设备。与底板 103 的内部插槽相比，在介质托架底板 107 上的设备安装称为“无遮蔽的”或“可访问的”。介质托架底板 107 有与之相关的控制器 107a。

在图 1 的例子中，该介质托架底板 107 还包括一个扩展器 107b。在 SAS 的术语集中，“扩展器”是提供带有到额外目标访问的发起者的

设备（反之亦然）。扩展器提供的功能与集线器或交换机提供的类似。扩展器可以连接到其他扩展器、SATA 目标设备、SAS 目标设备或 SAS 发起者。扩展器将发起者连接到目标设备，从而形成一个 SAS 域。

在系统 100 中，扩展器 107b 用于提高可以连接到控制器 105 的驱动器的数量。扩展器 107b 遵从在背景技术部分所讨论的 SAS 标准，该标准是定义扩展器 107b 如何与发起者和目标设备结合的接口标准。

在图 1 的例子中，扩展器 107b 是一个 6 端口 (x6) 扩展器。这是一个“边缘扩展器”。

如上所述，驱动器控制器 105 包含 SAS 控制器 106，后者通过边带信号连接到底板控制器 103a 和介质托架控制器 107a。

SAS 控制器 105 通过一号端口直接连接到底板 103 的驱动器 1—4。SAS 控制器 105 通过二号端口直接连接到驱动器 9、10 以及扩展器 107b。其次，扩展器 107b 连接到驱动器 5—8。这个配置只用于示范；这里阐明的概念可以扩展以便包括多个底板，不同的拓扑和驱动器数量。

数据存储控制器 105 的一个功能是扩展与底板 103 上的驱动器插槽相关联的状态 LED。为这个目的，控制器 105 的 RAID 软件使用传递给该底板控制器 103a 的 SES 或 SAF-TE 指令。这要求控制器 105 有特定的方法以便将它的端口（以及连接到这些端口的驱动器）映射到底板 103 的插槽上。

在有热插拔底板、但缺少 SAS 扩展器 107b 的服务器上，遵从并行 SCSI 标准的硬磁盘驱动器很容易与他们所驻留的各个驱动器插槽相关联。这是因为传统的并行 SCSI 底板设计为基于该驱动器所插入的插槽，将该 SCSI 目标 ID 设置为与该底板的插槽编号一样的驱动器 ID。这通过在底板连接器上用于每个插槽的跳线 (jumping) 和接地 (grounding) 信号实现。

图 2 说明了 I/O 控制器端口如何映射到如系统 100、但缺少扩展器的系统的驱动器插槽上。图 2 中，控制器 105 的每个端口与单个“物理口”(phy) 相关。这与以下说明的图 3 中的系统形成对照，图 3 中控制器端口可以是由使用超过一个物理口而建立一条宽链接 (wide link) 的宽端口 (wide port)。

图 2 中，每个物理口通过底板 103 直接连接到每个驱动器 101。驱动器与他们的插槽之间的关联允许由控制器轻松地控制每个驱动器插槽的状态 LED。端口 0 连接到插槽 1，以此类推。该控制器基于所指定的设备 ID 更新驱动器的状态，设备 ID 也对应于适当的驱动器插槽编号。

由此，参照图 1 和图 2，在不存在扩展器 107b 的情况下，将服务器插槽映射到驻留在这些插槽的驱动器上的逻辑方法，是使用该驱动器所连接的控制器 106 的端口（物理口）编号。这样允许控制器 106 基于端口（物理口）编号来控制这些插槽的状态 LED。

可是，在图 1 的配置中，实际上有扩展器 107b，以上所述的关联方法并不起作用。这是由于所有的 SAS 设备（包括扩展器 107b）有一个 64 位的全球范围名称（WWN）形式的地址，该地址在制造该设备的时候设置。在同扩展器和终端设备的通信期间，使用这个 SAS 地址。

图 3 说明用于定义对于图 1 中的系统而言，存储控制器 105 如何发现磁盘驱动器和指定设备 ID 的方法。图 3 的阴影部分代表通过介质托架底板 107 和扩展器 107b 而相连接的驱动器。设备 ID 是指定给驱动器 101 的编号，用于同该驱动器的通信并将该驱动器与一个服务器插槽相关联。

系统的布局配置为实现任意设备上的最低编号的插槽连接到用于该设备的最低编号的物理口。按照发现的先后次序来指定设备 ID。

不论服务器 100 是否有扩展器，物理口到插槽的布局都是正确的。当一个物理口上没有驱动器时，该设备 ID 保留。这样，如果稍后插入一个驱动器，编号次序将得以保存。

为发现驱动器 101，控制器 105 从包含物理口 0 的端口开始，遍历它的端口。该遍历流程在转移到下一个端口之前，先遍历物理口 0 上的所有设备，依此类推。当一个物理口不是一个宽端口的一部分时，该端口与该物理口是一样的。当多个物理口是一个宽端口的一部分时，通过这些物理口执行发现流程。

在图 3 的例子中，从用于位于底板 103 上的驱动器的设备 ID 0 开始，对应的设备 ID 与该插槽编号相关联。当没有驱动器存在于加到服务器插槽的物理口（属于该 I/O 控制器的，或是属于一个扩展器的），

跳过该设备 ID。这样，如果没有设备附加到对应于插槽 4 的物理口 3，就跳过该设备 ID。如果有设备存在于插槽 4，该设备将具有设备 ID 3。以这样的方式，指定设备 ID 0—3 给驱动器 1—4。

当一个扩展器连接到物理口（或者，在宽端口情况下是多个物理口），控制器 105 按照次序（从物理口 0 到物理口 x）发现该扩展器上的驱动器，并且以同样的方式指定设备 ID，为连接到空余插槽的物理口保留设备 ID，为现存的驱动器指定设备 ID。

在图 1 的配置并如图 3 所示，控制器 105 的物理口 4 和 5 配置为一个 2X 宽端口。他们扩展到底板 103 的插槽 5—8，因此这些插槽上的驱动器就具有这些设备 ID。物理口 6 和 7 连接到该底板的驱动器 9 和 10 上，这些驱动器就被指定相应的设备 ID。

可以用 SAS 等价术语来复述以上所述方法。在 SAS 系统中，对所有的设备到设备通信而言，标识序列是其关键。在标识序列中，位于一条运行链路上的 SAS 设备相互交换标识（IDENTIFY）地址帧。标识地址帧包含关于所连接 SAS 设备的信息，例如它的 SAS 地址，它是否支持特定协议（SMP, STP, SSP），它是目标设备，是发起者或两者都是，以及设备类型。设备类型可以是扩展器或是终端设备类型。当 SAS 发起者发现在这个 SAS 域中存在有什么设备，是否目标设备或发起者设备，以及他们支持的协议的时候，SAS 发起者使用标识地址帧中的信息。使用接着该 SAS 发现算法之后的 SMP 发现指令，将该信息回复给这个 SAS 发起者。通过 SMP 指令，发起者可以发现有关该扩展器的信息，例如该扩展器中的物理口数量，连接到扩展器特定物理口的设备类型，关于该扩展器制造商的信息，等等。

为将它的端口映射到驱动器插槽，控制器 105 以如下所述方式使用标识地址帧数据。从物理口 0 开始，如果收到标识地址帧数据，它会从中读取附加设备类型（ADT）。如果没有收到标识地址帧数据，即不存在设备，该 ADT 被内部设置为“没有设备”，然后控制器 105 将保留设备 ID 0。如果该 ADT 设置为“终端设备”，它指定设备 ID 0 给连接到对应于插槽 0 的物理口 0 的驱动器。

如果 ADT 设置为“边缘扩展器设备”或“输出扩展器设备”，控制器 105 通过将每个存在的设备的物理口指定为下一个设备 ID，以及

如果不存在驱动器则保留下一个设备 ID，来发现附加到该扩展器的所有设备。一旦该扩展器上的所有设备都被发现，该控制器 105 将在它的下一个物理口上继续以上映射流程。

当完成发现后，控制器 105 将拥有稀疏分布的设备 ID 列表。RAID 控制器 106 现在可以使用这些设备 ID 来操作插槽 LED。

图 4—图 6 说明了如以上所述的同样概念。图 4 中，类似系统 100，系统 400 是配置为 SAS 的存储系统。系统 400 中对应于系统 100 中类似设备的设备采用类似编号。可是，系统 400 中，扩展器 407b 有 12 个端口。图 5 说明了不存在该扩展器情况下的磁盘驱动器映射。图 6 中，4 个控制器端口配置为一个扩展到驱动器插槽 5—10 的 4X 宽端口，因此这些插槽中的驱动器被指定为对应的设备 ID 减 1。

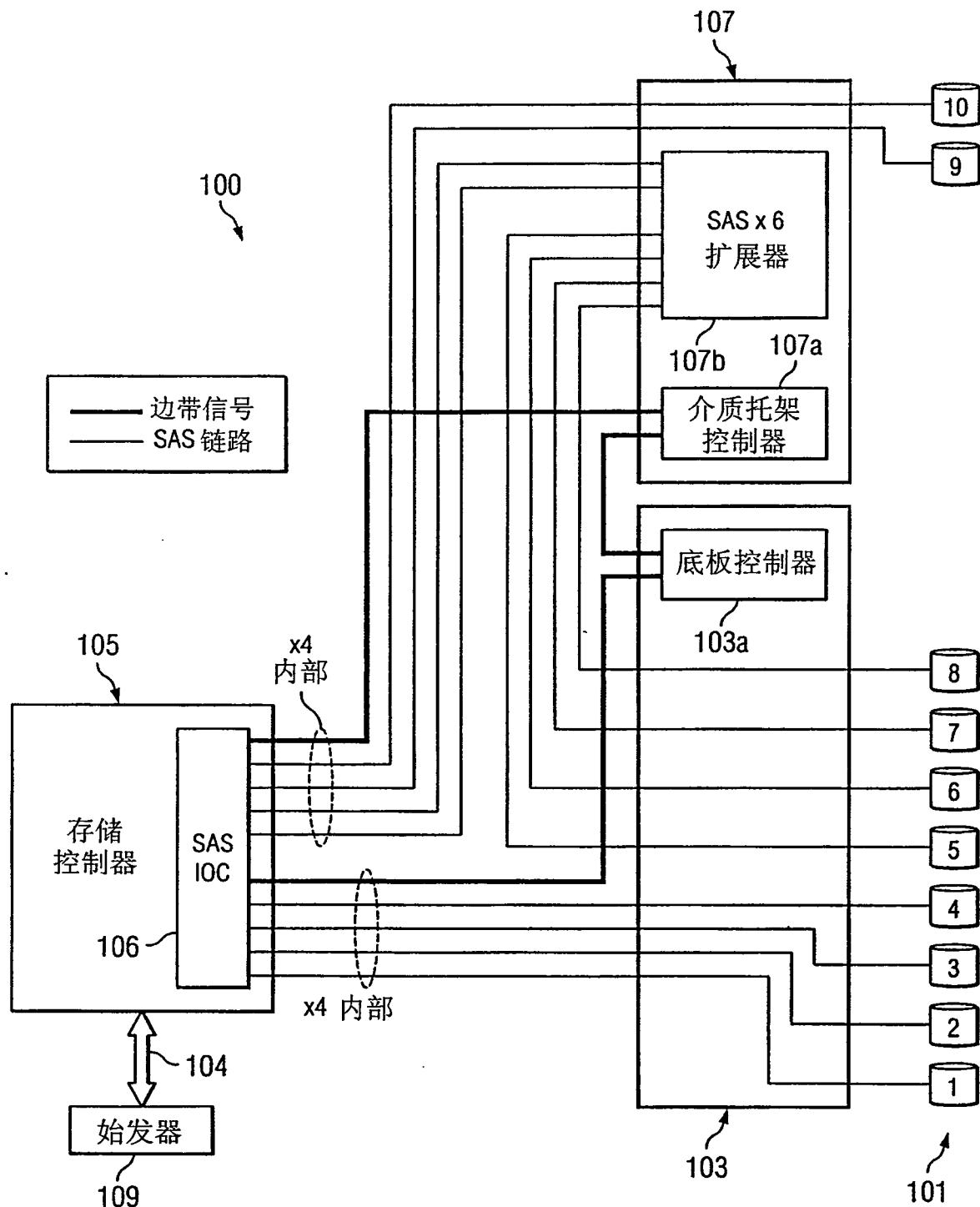


图1

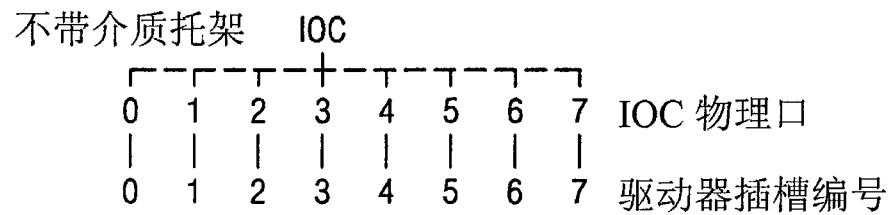


图2

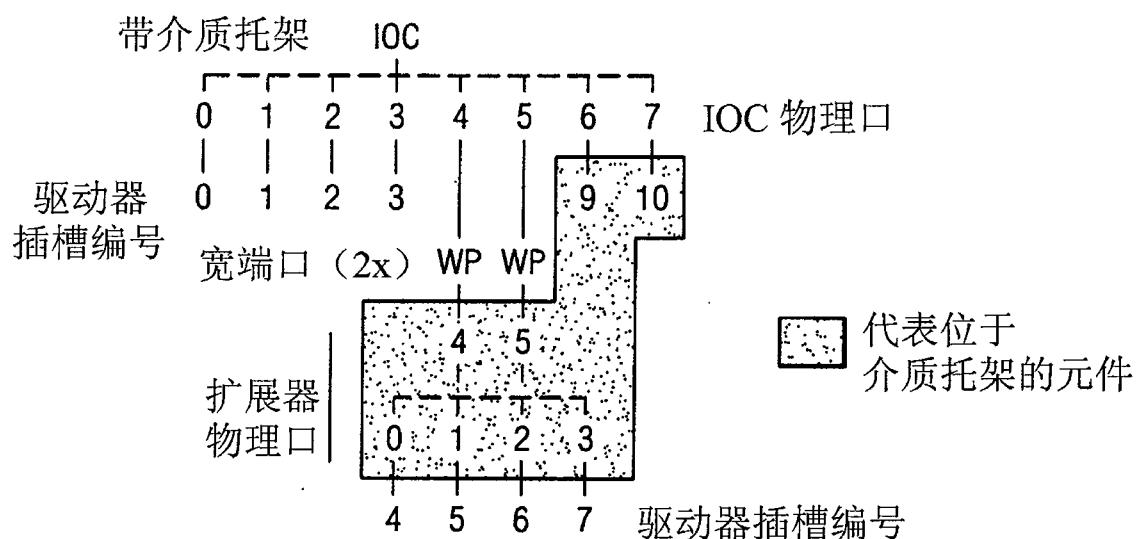
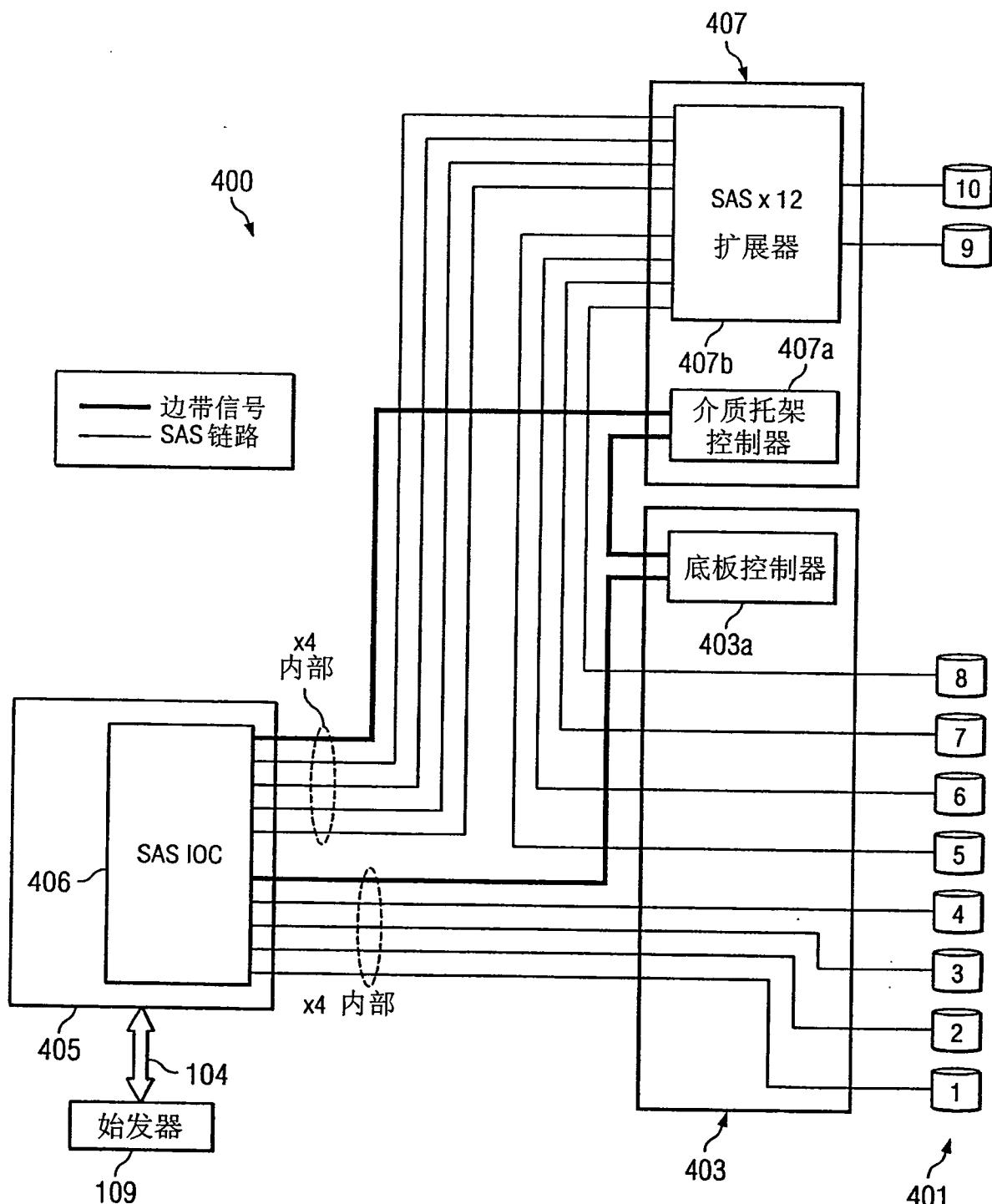


图3



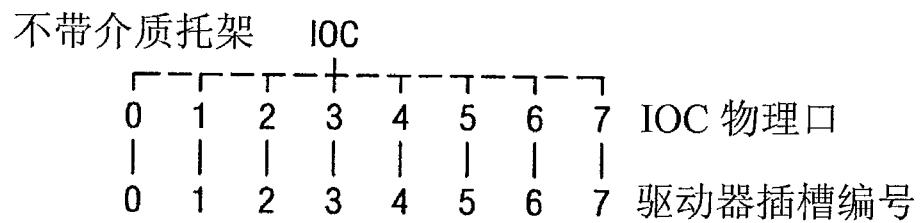


图5

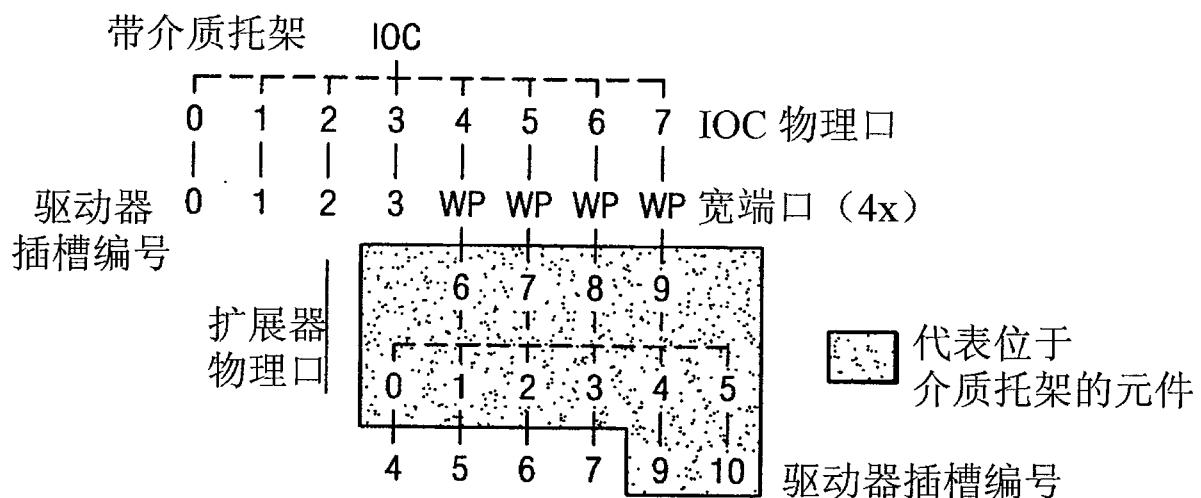


图6