



# [12] 发明专利申请公布说明书

[21] 申请号 200710005565.9

[43] 公开日 2007年9月26日

[11] 公开号 CN 101042632A

[22] 申请日 2007.2.12  
 [21] 申请号 200710005565.9  
 [30] 优先权  
     [32] 2006.3.20 [33] JP [31] 2006-076505  
 [71] 申请人 株式会社日立制作所  
     地址 日本东京都  
 [72] 发明人 中岛昭夫 八木泽育哉

[74] 专利代理机构 北京银龙知识产权代理有限公司  
 代理人 许静

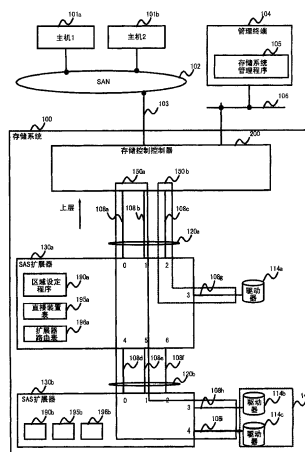
权利要求书 7 页 说明书 28 页 附图 22 页

## [54] 发明名称

控制物理链接分配的计算机系统及方法

## [57] 摘要

本发明提供一种控制物理链接分配的计算机系统及方法，即使对多个 SAS 目标装置中的特定 SAS 目标装置的访问集中，也尽量使计算机系统的吞吐量不下降。该计算机系统具备：多个 SAS 目标装置 (141)；SAS 启动器装置 (200)；以及交换机部 (130a、130b)，所述交换机部通过作为物理线路的物理链接与各 SAS 目标装置相连，且通过由多个物理链接而构成的宽链接与 SAS 启动器装置相连。SAS 启动器装置对将宽链接中的哪个物理链接分配给哪个 SAS 目标装置进行控制，由此，经由分配给该 SAS 目标装置的物理链接进行从 SAS 启动器装置向 SAS 目标装置的访问，而不经由未分配给该 SAS 目标装置的物理链接进行该访问。



1. 一种计算机系统，其特征在于，具备：

多个 SAS 目标装置，所述 SAS 目标装置是遵从 SAS 协议的目标；

SAS 启动器装置，所述 SAS 启动器装置是遵从 SAS 协议的启动器；以及

交换机部，其通过作为物理线路的物理链接与各 SAS 目标装置相连，并且通过由多个物理链接构成的宽链接与所述 SAS 启动器装置相连，

所述交换机部是一个或多个交换机装置，当该交换机部为多个交换机装置时，以与所述 SAS 启动器装置相连的交换机装置为起始级联连接，一个交换机装置与另一交换机装置彼此通过宽链接相连，

所述计算机系统还具备物理链接分配控制部，

所述物理链接分配控制部，控制将所述交换机部和所述 SAS 启动器装置之间的宽链接以及所述交换机部内的宽链接中至少一个宽链接中的哪条物理链接分配给哪个 SAS 目标装置，

从所述 SAS 启动器装置向 SAS 目标装置的 I/O，经由分配给该 SAS 目标装置的物理链接进行，无法经由未分配给该 SAS 目标装置的物理链接进行。

2. 根据权利要求 1 所述的计算机系统，其特征在于，

还具有监视部，其关于各 SAS 目标装置，对作为 SAS 目标装置所请求的 I/O 性能的 I/O 请求性能进行监视，

所述物理链接分配控制部，根据对各 SAS 目标装置的 I/O 请求性能，控制分配给各 SAS 目标装置的物理链接的数量。

3. 根据权利要求 2 所述的计算机系统，其特征在于，

所述物理链接分配控制部，使分配给 I/O 请求性能高的 SAS 目标装置的物理链接数多于分配给 I/O 请求性能低的 SAS 目标装置的物理链接数。

4. 根据权利要求 1 所述的计算机系统，其特征在于，

在所述多个 SAS 目标装置中，通信协议不同的 SAS 目标装置并存，

所述物理链接分配控制部，根据 SAS 目标装置的通信协议的种类，控制分配给各 SAS 目标装置的物理链接数。

5. 根据权利要求4所述的计算机系统，其特征在于，

所述物理链接分配控制部，使分配给以转发效率低的通信协议进行通信的SAS目标装置的物理链接数，多于分配给以转发效率高的通信协议通信的SAS目标装置的物理链接数。

6. 根据权利要求4所述的计算机系统，其特征在于，

所述物理链接分配控制部，使分配给以转发效率高的通信协议进行通信的SAS目标装置的物理链接数，多于分配给以转发效率低的通信协议通信的SAS目标装置的物理链接数。

7. 根据权利要求1所述的计算机系统，其特征在于，

各SAS目标装置是存储装置，

两个以上存储装置通过按照RAID规则分组而成为RAID组，

所述物理链接分配控制部以RAID组单位控制分配的物理链接数。

8. 根据权利要求1所述的计算机系统，其特征在于，

还具备区域设定部，其进行将从所述SAS启动器装置至所述多个SAS目标装置的访问路径分配给多个逻辑区域的区域设定，

所述物理链接分配控制部，根据所述被设定的多个区域，控制分配给各SAS目标装置的物理链接数。

9. 根据权利要求8所述的计算机系统，其特征在于，

所述物理链接分配控制部，当多个SAS目标装置属于一个区域时，在所述一个区域的范围内控制分配给所述多个SAS目标装置的每一个的物理链接数。

10. 根据权利要求8所述的计算机系统，其特征在于，

还具备资源监视部，其定期发送资源监视命令，该资源监视命令用于定期监视连接多个SAS目标装置的机架的资源，

所述区域设定部，设定在所述资源监视命令的发送中不使用而在向SAS目标装置的I/O中使用的I/O专用区域、和在所述I/O中不使用而在所述资源监视命令的发送中使用的资源监视专用区域。

11. 根据权利要求2所述的计算机系统，其特征在于，

所述SAS启动器装置，从存在于所述计算机系统外部的访问源接收I/O

请求，对遵从该 I/O 请求的第一 SAS 目标装置进行第一 I/O，并且与来自所述访问源的 I/O 请求无关地在后端对第二 SAS 目标装置进行第二 I/O，

所述物理链接分配控制部，根据所述第一 I/O 的 I/O 请求性能和所述第二 I/O 的 I/O 请求性能，控制分配给所述第一 SAS 目标装置的物理链接数以及分配给所述第二 SAS 目标装置的物理链接数。

12. 根据权利要求 11 所述的计算机系统，其特征在于，

当所述第一 I/O 的 I/O 请求性能比所述第二 I/O 的 I/O 请求性能高时，所述物理链接分配控制部使分配给所述第一 SAS 目标装置的物理链接数多于分配给所述第二 SAS 目标装置的物理链接数。

13. 根据权利要求 1 所述的计算机系统，其特征在于，

所述 SAS 启动器装置，分别从存在于所述计算机系统外部的多个访问源接收 I/O 请求，对遵从该 I/O 请求的 SAS 目标装置进行 I/O，

所述物理链接分配控制部，根据所述多个访问源各自的优先度，使分配给遵从来自优先度高的访问源的 I/O 请求的 SAS 目标装置的物理链接数，多于分配给遵从来自优先度低的访问源的 I/O 请求的 SAS 目标装置的物理链接数。

14. 根据权利要求 1 所述的计算机系统，其特征在于，

所述物理链接分配控制部，在所述交换机部与所述 SAS 启动器装置之间的宽链接以及所述交换机部内的宽链接中的至少一个宽链接中，使分配给某 SAS 目标装置的物理链接数为零，或者解除将所述某 SAS 目标装置与所述交换机部相连的物理链接的分配，由此防止对所述某 SAS 目标装置的 I/O。

15. 根据权利要求 1 所述的计算机系统，其特征在于，

所述一个或多个交换机装置各自具备：与多个物理链接分别相连的多个通信端口；以及

存储交换机控制信息的存储区，

所述交换机控制信息中，关于所述多个通信端口的每一个，记录有表示不经由其他通信端口而直接连接的装置的直接装置数据、和表示经由其他通信端口间接连接的 SAS 目标装置的间接装置数据，

当关于通信端口的所述直接装置数据是表示 SAS 目标装置的数据时，该

SAS 目标装置是分配了与该通信端口相连的物理链接的 SAS 目标装置，

关于通信端口的所述间接装置数据是表示分配了与该通信端口相连的物理链接的 SAS 目标装置的数据，

所述物理链接分配控制部，通过更新所述一个或多个交换机装置中的至少一个交换机装置的所述交换机控制信息，控制分配给 SAS 目标装置的物理链接。

16. 根据权利要求 15 所述的计算机系统，其特征在于，

所述一个或多个交换机的每一个，若经过未分配给某 SAS 目标装置的物理链接接收到用于确立与所述某 SAS 目标装置的连接的连接打开指示，则经过该物理链接向该连接打开指示的发送源返回错误，不经过该物理链接确立连接。

17. 根据权利要求 1 所述的计算机系统，其特征在于，

在所述计算机系统中，可通信地连接有用于管理所述计算机系统的管理终端，

所述物理链接分配控制部，按照来自所述管理终端的指示，控制将哪条物理链接分配给哪个 SAS 目标装置。

18. 根据权利要求 1 所述的计算机系统，其特征在于，

所述计算机系统是具备多个存储装置的存储系统，

所述多个 SAS 目标装置是所述多个存储装置，

所述 SAS 启动器装置是控制向各存储装置的 I/O 的控制器，

所述交换机部是多个交换机装置，

所述多个交换机装置各自具有：与多个物理链接分别连接的多个通信端口、以及存储了交换机控制信息的存储区，

在所述交换机控制信息中，关于所述多个通信端口的每一个，记录有：表示不经过其他通信端口而直接连接的装置的直接装置数据、以及表示经过其他通信端口间接连接的存储装置的间接装置数据，

当关于通信端口的所述直接装置数据是表示存储装置的数据时，该存储装置是分配了与该通信端口相连的物理链接的存储装置，

关于通信端口的所述间接装置数据，是表示分配了与该通信端口相连的

物理链接的存储装置的数据，

所述存储系统还具备监视部，该监视部关于各存储装置，监视作为存储装置所请求的 I/O 性能的 I/O 请求性能，

所述物理链接分配控制部，根据针对各存储装置的 I/O 请求性能，更新所述多个交换机装置中的至少一个交换机装置的所述交换机控制信息，由此控制分配给各存储装置的物理链接数。

19. 根据权利要求 18 所述的计算机系统，其特征在于，

两个以上的存储装置通过按照 RAID 规则分组而形成 RAID 组，

所述物理链接分配控制部以 RAID 组单位来控制分配的物理链接数。

20. 根据权利要求 1 所述的计算机系统，其特征在于，

所述计算机系统是具备多个存储装置的存储系统，

所述多个 SAS 目标装置是所述多个存储装置，

在所述多个存储装置中，作为进行遵从 SCSI 协议的通信的存储装置的 SCSI 存储装置、与作为进行遵从 ATA 协议的通信的存储装置的 ATA 存储装置并存，

所述 SAS 启动器装置是控制向各存储装置的 I/O 的控制器，

所述交换机部是多个交换机装置，

所述多个交换机装置各自具有：与多个物理链接分别相连的多个通信端口、以及存储了交换机控制信息的存储区，

在所述交换机控制信息中，关于所述多个通信端口的每一个，记录有：表示不经过其他通信端口而直接连接的装置的直接装置数据、以及表示经过其他通信端口间接连接的存储装置的间接装置数据，

当关于通信端口的所述直接装置数据是表示存储装置的数据时，该存储装置是分配了与该通信端口相连的物理链接的存储装置，

关于通信端口的所述间接装置数据是表示分配了与该通信端口相连的物理链接的存储装置的数据，

所述物理链接分配控制部，根据各存储装置的通信协议是 SCSI 还是 ATA，更新所述多个交换机装置中至少一个交换机装置的所述交换机控制信息，由此控制分配给各存储装置的物理链接数。

21. 一种存储系统，其特征在于，具备：

多个存储装置，所述存储装置是遵从 SAS 协议的目标；

控制器，所述控制器是遵从 SAS 协议的启动器，控制向各存储装置的 I/O；  
以及

交换机部，所述交换机部通过作为物理线路的物理链接与各存储装置相连，并且通过作为由多个物理链接构成的逻辑路径的宽链接与所述控制器相连，

两个以上的存储装置，通过遵从 RAID 规则分组而形成 RAID 组，

所述交换机部是多个交换机装置，以与所述控制器相连的交换机装置为起始级联连接，一个交换机装置与另一交换机装置互相通过宽链接相连，

所述多个交换机装置各自具有：与多个物理链接分别相连的多个通信端口、以及存储了交换机控制信息的存储区，

在所述交换机控制信息中，关于所述多个通信端口的每一个，记录有：表示不经过其他通信端口而直接连接的装置的直接装置数据、以及表示经过其他通信端口间接连接的存储装置的间接装置数据，

当关于通信端口的所述直接装置数据是表示存储装置的数据时，该存储装置是分配了与该通信端口相连的物理链接的存储装置，

关于通信端口的所述间接装置数据是表示分配与该通信端口相连的物理链接的存储装置的数据，

所述控制器，关于各 RAID 组监视作为 RAID 组所请求的 I/O 性能的 I/O 请求性能，根据该监视结果，更新所述多个交换机装置中至少一个交换机装置的所述交换机控制信息，由此控制将所述交换机控制部与所述控制器之间的宽链接以及所述交换机部内的宽链接中的至少一个宽链接中的哪条物理链接分配给哪个 RAID 组，

从所述控制器向存储装置的 I/O，经由分配给该存储装置的物理链接进行，而无法经由未分配给该存储装置的物理链接进行。

22. 一种存储系统，其特征在于，具备：

多个存储装置，所述存储装置是遵从 SAS 协议的目标；

控制器，所述控制器是遵从 SAS 协议的启动器，控制向各存储装置的 I/O；

以及

交换机部，所述交换机部通过作为物理线路的物理链接与各存储装置相连，并且通过作为由多个物理链接构成的逻辑路径的宽链接与所述控制器相连，

在所述多个存储装置中，作为进行遵从 SCSI 协议的通信的存储装置的 SCSI 存储装置、与作为进行遵从 ATA 协议的通信的存储装置的 ATA 存储装置并存，

所述交换机部是多个交换机装置，以与所述控制器相连的交换机装置为起始级联连接，一个交换机装置与另一交换机装置通过宽链接互相连接，

所述多个交换机装置各自具有：与多个物理链接分别相连的多个通信端口、以及存储了交换机控制信息的存储区，

在所述交换机控制信息中，关于所述多个通信端口的每一个，记录有：表示不经过其他通信端口而直接连接的装置的直接装置数据、以及表示经过其他通信端口间接连接的存储装置的间接装置数据，

当关于通信端口的所述直接装置数据是表示存储装置的数据时，该存储装置是分配了与该通信端口相连的物理链接的存储装置，

关于通信端口的所述间接装置数据是表示分配与该通信端口相连的物理链接的存储装置的数据，

所述控制器，根据各存储装置的通信协议是 SCSI 还是 ATA，来更新所述多个交换机装置中至少一个交换机装置的所述交换机控制信息，由此控制将所述交换机控制部与所述控制器之间的宽链接以及所述交换机部内的宽链接中的至少一个宽链接中的哪条物理链接分配给哪个存储装置，

从所述控制器向存储装置的 I/O，经由分配给该存储装置的物理链接进行，而无法经由未分配给该存储装置的物理链接来进行。

23. 根据权利要求 1 所述的计算机系统，其特征在于，

所述被分配控制的物理链接是虚拟物理链接，

所述虚拟物理链接是通过对一个物理链接分时而形成的逻辑链接，

从所述 SAS 启动器装置向 SAS 目标装置的 I/O，经由分配给该 SAS 目标装置的逻辑链接进行，而无法经由未分配给该 SAS 目标装置的逻辑链接进行。



## 控制物理链接分配的计算机系统及方法

### 技术领域

本发明涉及具备物理链接的计算机系统，具体而言，涉及例如在后端（back-end）具有用于进行遵从 SAS（Serial Attached SCSI）的通信的物理链接的计算机系统。

### 背景技术

作为计算机系统的一种，已知具备多个存储装置（例如硬盘驱动器或 DVD（Digital Versatile Disk）驱动器等盘驱动器）的存储系统。存储系统可以从访问源装置（例如主机或其他存储系统）接收 I/O 命令（作为写命令或读命令的输入/输出命令），将遵从所接收到的 I/O 命令的数据写入存储装置，或者将其从存储装置读出并发送至访问源装置。

作为这种存储系统，已知例如特开 2005-149173 号公报中公开的盘装置。该盘装置中，在光纤通道仲裁环路（Fiber Channel-Arbitrated Loop（FC-AL））上连接了多个盘驱动器。

【专利文献 1】特开 2005-149173 号公报

### 发明内容

作为通信 I/F，已知有 SAS（Serial Attached SCSI）。SAS 是可以利用相同物理线路来连接不同协议种类（例如 SCSI 和 ATA）的盘驱动器的 I/F 技术。作为可以在后端进行遵从 SAS 的通信的计算机系统，考虑到具备多个盘驱动器和控制对各盘驱动器的访问的控制器的存储系统。在这种情况下，在该存储系统中，考虑到作为用于扩展连接盘驱动器的交换机而配备 SAS 扩展器（expander），该 SAS 扩展器具备多个物理通信端口（以下称为物理端口）。可以具有多个 SAS 扩展器，当控制器位于最上层时，可以从最上层至下层将多个 SAS 扩展器级联（cascade）连接（例如，以串联或树形结构连接）。但控制器未必是最上层，在 SAS 扩展器的任何位置都可以连接 SAS 启动器（initiator）。可以在各 SAS 扩展器的任意物理端口连接控制器、上层 SAS 扩

展器、下层 SAS 扩展器或盘驱动器。

在该存储系统中，在后端，在控制器与 SAS 扩展器之间、SAS 扩展器间、以及 SAS 扩展器与盘驱动器之间设有物理线路（以下称为物理链接）。控制器（SAS 启动器），当访问盘驱动器（SAS 目标）时可以从控制器至盘驱动器经过各物理链接来访问盘驱动器。具体而言，例如当控制器访问与控制器下方的第一 SAS 扩展器相连的盘驱动器时，可以经过连接控制器与第一 SAS 扩展器的第一物理链接、第一 SAS 扩展器、以及连接第一 SAS 扩展器与盘驱动器的第二物理链接来访问盘驱动器。

在 SAS 中存在可以将一个装置间并列的多个物理链接汇集为一个逻辑的链接（以下称为逻辑链接）的、被称为“宽链接”的技术。即宽链接可以由例如在 SAS 扩展器~SAS 扩展器间、以及 SAS 启动器~SAS 扩展器间连接一个装置与另一装置的并列的多个物理链接形成。另外，也可以没有 SAS 启动器的指示，在 SAS 扩展器~SAS 扩展器间、或 SAS 启动器~SAS 扩展器间自动形成宽链接。

另外，SAS 中有如下技术：可以通过将一个物理链接分时来形成多个虚拟物理链接（以后称为逻辑链接）。从而可以在形成宽链接的各个物理链接中构成多个逻辑链接。例如，当以转发速度为 6Gbps 的两条物理链接形成一条宽链接时，由于可以在一个物理链接中形成转发速度为其 1/4 即 1.5Gbps 的逻辑链接，因此可以在该宽链接中形成共计 8 个逻辑链接。在这种情况下，SAS 启动器装置可以同时与 8 个 SAS 目标装置确立连接，同时对帧的收发进行转发。

控制器经由构成宽链接的多个物理链接可同时与多个装置确立连接。当控制器与 SAS 扩展器间通过宽链接相连、且在该 SAS 扩展器上连接有多个盘驱动器时，控制器能够以构成宽链接的物理链接数同时对多个盘驱动器确立连接。1 个控制器通过宽链接可以同时与多个盘驱动器执行帧转发。

在该存储系统中，在例如以下第一和第二种情况中，分别会产生该情况特有的问题。

所谓第一种情况，是在 SAS 扩展器上连接了通信 I/F 不同的盘驱动器的情况。例如，是在一个或多个 SAS 扩展器上连接了具有 SAS I/F 的盘驱动器

(以下称为 SAS 驱动器)和具有 SATA (Serial ATA) I/F 的盘驱动器(以下称为 SATA 驱动器)的情况。通过使用相同物理链接转发 STP (Serial ATA Tunneled Protocol) 帧可以实现该情况,所述 STP 帧将用于访问 SAS 驱动器或 SAS 装置的 SSP (Serial SCSI Protocol) 帧与用于访问 SATA 驱动器的 SATA 协议隧通 (tunnelling)。在此, SATA 与 SAS 中的转发效率不同(具体而言, SATA 的转发效率比 SAS 的转发效率低,物理链接的占有时间长)。因此,当 SATA 帧的 I/O 转发请求较多时,可能给针对 SAS 驱动器的吞吐量 (throughput) 造成不良影响。此类问题不仅限于这两种 I/F,也存在于其他种类的 I/F 间。

所谓第二种情况,是使两个以上盘驱动器成为遵从 RAID (Redundant Array of Independent(or Inexpensive)Disks)规则的一个组(以下称为 RAID 组),将 RAID 组与 SAS 扩展器相连的情况。由于多个 SAS 扩展器级联连接,因此有时从控制器至某个 RAID 组的路径中的宽链接、和从控制器至其他 RAID 组的路径中的宽链接在后端内的上层相同。所以,当对特定 RAID 组的访问增加时,可能会给针对其他 RAID 组的吞吐量造成不良影响。

以上问题不限于存储系统,也会存在于其他种类的计算机系统中。例如,在计算机系统中存在多个 SAS 目标装置的情况下,当对特定 SAS 目标装置的访问集中时,该计算机系统的吞吐量可能会下降。

另外,由于 SAS 协议的制约,无法由 SAS 启动器装置控制从构成宽链接的多个物理链接中选择执行帧转发的物理链接的步骤。

因此,本发明的目的在于,即使对多个 SAS 目标装置中的特定 SAS 目标装置的访问集中,也尽量使计算机系统的吞吐量不下降。理想的是不违反后端的通信协议地实现该目的。

从后述说明中可以了解本发明的其他目的。

按照本发明的计算机系统具备:多个 SAS 目标装置,是遵从 SAS 协议的目标; SAS 启动器装置,是遵从 SAS 协议的启动器;交换机部,其通过作为物理线路的物理链接与各 SAS 目标装置相连,且通过由多个物理链接构成的宽链接与所述 SAS 启动器装置相连。所述交换机部是一个或多个交换机装置,当该交换机部为多个交换机装置时,以与所述 SAS 启动器装置相连的交换机

装置为起始级联连接，一交换机装置与另一交换机装置通过宽链接互相连接。所述计算机系统还具备物理链接分配控制部。所述物理链接分配控制部，控制将所述交换机部与所述 SAS 启动器装置之间的宽链接以及所述交换机部内的宽链接中至少一个宽链接中的哪个物理链接分配给哪个 SAS 目标装置。从所述 SAS 启动器装置向 SAS 目标装置的 I/O，经由分配给该 SAS 目标装置的物理链接来进行，无法经由未分配给该物理目标装置的物理链接进行。

在第一实施方式中可以进一步具有监视部，其关于各 SAS 目标装置，监视作为 SAS 目标装置所请求的 I/O 性能的 I/O 请求性能。所述物理链接分配控制部可以根据对应于各 SAS 目标装置的 I/O 请求性能，控制分配给各 SAS 目标装置的物理链接数。

在第二实施方式中，第一实施方式中，所述物理链接分配控制部，可以使分配给 I/O 请求性能高的 SAS 目标装置的物理链路数多于分配给 I/O 请求性能低的 SAS 目标装置的物理链接数。

在第三实施方式中，所述多个 SAS 目标装置中并存有通信协议不同的 SAS 目标装置。所述物理链接分配控制部可以根据 SAS 目标装置的通信协议种类，控制分配给各 SAS 目标装置的物理链接数。

在第四实施方式中，所述第三实施方式中，所述物理链接分配控制部，可以使分配给以转发效率低的通信协议通信的 SAS 目标装置的物理链接数多于分配给以转发效率高的通信协议通信的 SAS 目标装置的物理链接数。反之，可以使分配给以转发效率高的通信协议通信的 SAS 目标装置的物理链接数多于分配给以转发效率低的通信协议通信的 SAS 目标装置的物理链接数。

在第五实施方式中，各 SAS 目标装置为存储装置。可以通过将两个以上存储装置按照 RAID 规则分组而成为 RAID 组。所述物理链接分配控制部可以以 RAID 组单位来控制分配物理链接数。

在第六实施方式中可以进一步具有区域 (zone) 设定部，其进行区域设定，将从所述 SAS 启动器装置至所述多个 SAS 目标装置的访问路径分配到多个逻辑区域。所述物理链接分配控制部可以根据所述设定的多个区域来控制分配给各 SAS 目标装置的物理链接数。

在第七实施方式中，所述第六实施方式中，当多个 SAS 目标装置属于一

个区域时，所述物理链接分配控制部可以在所述一个区域的范围内控制分配给所述多个 SAS 目标装置的每个的物理链接数。

在第八实施方式中可以进一步具有资源监视部，其定期发送资源监视命令，该资源监视命令用于定期监视连接了多个 SAS 目标装置的机架的资源（例如 SAS 目标装置）。在这种情况下，所述区域设定部可以设定 I/O 专用区域和资源监视专用区域，在所述资源监视命令的发送中不使用、而在对 SAS 目标装置的 I/O 中使用所述 I/O 专用区域，在所述 I/O 中不使用、而在所述资源监视命令的发送中使用所述资源监视专用区域。

在第九实施方式中，所述第一实施方式中，所述 SAS 启动器装置可以从存在于所述计算机系统外部的访问源接收 I/O 请求，对遵从该 I/O 请求的第一 SAS 目标装置进行第一 I/O，且与来自所述访问源的 I/O 请求无关地、在后端对第二 SAS 目标装置进行第二 I/O。所述物理链接分配控制部根据所述第一 I/O 的 I/O 请求性能和所述第二 I/O 的 I/O 请求性能，控制分配给所述第一 SAS 目标装置的物理链接数和分配给所述第二 SAS 目标装置的物理链接数。

在第十实施方式中，所述第九实施方式中，当所述第一 I/O 的 I/O 请求性能高于所述第二 I/O 的 I/O 请求性能时，所述物理链接分配控制部可以使分配给所述第一 SAS 目标装置的物理链接数多于分配给所述第二 SAS 目标装置的物理链接数。

在第十一实施方式中，所述 SAS 启动器装置从存在于所述计算机系统外部的多个访问源分别接收 I/O 请求，对遵从该 I/O 请求的 SAS 目标装置进行 I/O。所述物理链接分配控制部可以根据所述多个访问源各自的优先度，使分配给遵从来自优先度高的访问源的 I/O 请求的 SAS 目标装置的物理链接数，多于分配给遵从来自优先度低的访问源的 I/O 请求的 SAS 目标装置的物理链接数。

在第十二实施方式中，所述物理链接分配控制部，在所述交换机部与所述 SAS 启动器装置之间的宽链接以及所述交换机部内的宽链接中至少一个宽链接中，使分配给某个 SAS 目标装置的物理链接数为零，或者解除连接所述 SAS 目标装置和所述交换机部的物理链接的分配，由此可以防止对所述某个 SAS 目标装置的 I/O。

在第十三实施方式中，所述一个或多个交换机装置各自可以具备：分别与多个物理链接相连的多个通信端口、以及存储了交换机控制信息的存储区。在所述交换机控制信息中，关于所述多个通信端口的每个，记录有表示不通过其他通信端口而直接连接的装置的直接装置数据、以及表示通过其他通信端口间接连接的 SAS 目标装置的间接装置数据。当关于通信端口的所述直接装置数据是表示 SAS 目标装置的数据时，该 SAS 目标装置可以是分配了与该通信端口连接的物理链接的 SAS 目标装置。关于通信端口的所述间接装置数据可以是表示分配了与该通信端口相连的物理链接的 SAS 目标装置的数据。所述物理链接分配控制部，通过对所述一个或多个交换机装置中至少一个交换机装置的所述交换机控制信息进行更新，控制分配给 SAS 目标装置的物理链接。

在第十四实施方式中，所述第十三实施方式中，所述一个或多个交换机的每个，若经过未分配给某 SAS 目标装置的物理链接接收到用于确立与所述某 SAS 目标装置的连接的连接打开（Connection Open）指示，则可以经过该物理链接向该连接打开指示的发送源返回错误，不经过该物理链接确立连接。

在第十五实施方式中，在所述计算机系统上也可以可通信地连接用于管理所述计算机系统的管理终端。所述物理链接分配控制部，可以按照来自所述管理终端的指示，控制将哪个物理链接分配给哪个 SAS 目标装置。

在第十六实施方式中，所述计算机系统可以是具有多个存储装置的存储系统。所述多个 SAS 目标装置可以是所述多个存储装置。所述 SAS 启动器装置可以是控制向各存储装置的 I/O 的控制器。所述交换机部可以是多个交换机装置。所述多个交换机装置各自可以具有与多个物理链接分别连接的多个通信端口、以及存储了交换机控制信息的存储区。在所述交换机控制信息中，关于所述多个通信端口的每个，可以记录有表示不经过其他通信端口而直接连接的装置的直接装置数据、以及表示经过其他通信端口间接连接的存储装置的间接装置数据。当关于通信端口的所述直接装置数据是表示存储装置的数据时，该存储装置可以是分配了与该通信端口相连的物理链接的存储装置。关于通信端口的所述间接装置数据，可以是表示分配了与该通信端口相连的物理链接的存储装置的数据。所述存储系统可以进一步具有监视部，该监视

部关于各存储装置，监视作为存储装置所请求的 I/O 性能的 I/O 请求性能。所述物理链接分配控制部，根据对应于各存储装置的 I/O 请求性能，更新所述多个交换机装置中至少一个交换机装置的所述交换机控制信息，由此可以控制分配给各存储装置的物理链接数。

在第十七实施方式中，所述第十六实施方式中，两个以上存储装置通过遵从 RAID 规则分组而成为 RAID 组，所述物理链接分配控制部可以以 RAID 组单位来控制分配的物理链接数。

在第十八实施方式中，所述计算机系统可以是具备多个存储装置的存储系统。所述多个 SAS 目标装置可以是所述多个存储装置。所述多个存储装置中，作为进行遵从 SCSI 协议的通信的存储装置的 SCSI 存储装置、作为进行遵从 ATA 协议的通信的存储装置的 ATA 存储装置可以并存。所述交换机部可以是多个交换机装置。所述多个交换机装置各自可以具有：与多个物理链接分别连接的多个通信端口、以及存储了交换机控制信息的存储区。所述交换机控制信息中，关于所述多个通信端口的每个，可以记录有表示不经过其他通信端口而直接连接的装置的直接装置数据、以及表示经过其他通信端口间接连接的存储装置的间接装置数据。当关于通信端口的所述直接装置数据是表示存储装置的数据时，该存储装置可以是分配了与该通信端口相连的物理链接的存储装置。关于通信端口的所述间接装置数据可以是表示分配了与该通信端口相连的物理链接的存储装置的数据。所述物理链接分配控制部，根据各存储装置的通信协议是 SCSI 还是 ATA，来更新所述多个交换机装置中至少一个交换机装置的所述交换机控制信息，由此可以控制分配给各存储装置的物理链接数。

在第十九实施方式中，所述被分配控制的物理链接可以是虚拟物理链接。所述虚拟物理链接是通过对一个物理链接进行分时而形成的逻辑链接。从所述 SAS 启动器装置向 SAS 目标装置的 I/O，经由分配给该 SAS 目标装置的逻辑链接进行，而无法经由未分配给该 SAS 目标装置的逻辑链接进行。所述 SAS 启动器装置，在一个物理链接中的逻辑链接中可以同时与多个所述 SAS 目标装置确立连接。

上述各部也可以称为各单元。可以通过硬件（例如电路）、计算机程序、

或它们的组合（例如，读取并执行计算机程序的一个或多个 CPU）来实现各部。可以从计算机系统所具备的存储资源（例如存储器）读入各计算机程序。在该存储资源中，可以通过 CD-ROM 或 DVD（Digital Versatile Disk）等记录介质进行安装，也可以通过因特网或 LAN 等通信网络进行下载。

另外，上述物理链接分配控制部和监视部等中的至少一个，可以安装在 SAS 启动器装置上。

根据本发明，即使向多个 SAS 目标装置中的特定 SAS 目标装置的访问集中，也可以尽量使计算机系统的吞吐量不下降。

#### 附图说明

图 1 是表示本发明的一种实施方式的存储系统的结构例的框图。

图 2 是表示存储控制控制器 200 的结构例的框图。

图 3 表示控制程序 300 的结构例。

图 4 表示 SAS 扩展器 130a 内的直接装置表（direct device table）195a 的结构例以及在该表 195a 中登录的信息。

图 5 表示 SAS 扩展器 130b 内的直接装置表 195b 的结构例以及在该表 195b 中登录的信息。

图 6 表示 SAS 扩展器 130a 内的扩展器路由表 196a 的结构例以及在该表 196a 中登录的信息。

图 7 表示 SAS 扩展器 130b 内的扩展器路由表 196b 的结构例以及在该表 196b 中登录的信息。

图 8 表示存储系统启动时进行的处理流程的一例。

图 9 表示变更分配给区域 150 的物理链接 108 的数量的处理流程的一例。

图 10 表示在图 9 的步骤 600 中进行的处理流程的一例。

图 11 表示区域状态管理表 700 的结构例。

图 12A 表示进行物理链接的分配变更的第一情况中的分配变更前的物理链接的分配。

图 12B 表示该分配变更后的物理链接的分配。

图 13 表示进行物理链接的分配变更的第二情况中的分配变更前的物理链接的分配。



图 14 表示该分配变更后的物理链接的分配。

图 15A 表示进行物理链接的分配变更的第三情况中的分配变更前的物理链接的分配。

图 15B 表示该分配变更后的物理链接的分配。

图 16 是本发明的一种实施方式的第一变形例的说明图。

图 17 是本发明的一种实施方式的第二变形例的说明图。

图 18 是本发明的一种实施方式的第三变形例的说明图。

图 19 是本发明的一种实施方式的第四变形例的说明图。

图 20 是本发明的一种实施方式的第五变形例的说明图。

图 21 表示用于在 SAS 目标装置与 SAS 启动器装置之间确立连接的处理流程的一例。

图 22 表示在一个宽链接中并存有物理链接和逻辑链接的例子。

#### 符号说明

100: 存储系统、101: 主机、102: SAN(存储区网络(storage area network))、103: 接口、104: 管理终端、105: 存储系统管理程序、106: 管理用网络、108: 物理链接、114: 盘驱动器、120: 宽链接、130: SAS 扩展器、140: RAID 组、150 区域、190: 区域设定程序、195: 直接装置表、196: 扩展器路由表、200: 存储控制控制器、210: 主机 I/F 控制器、211: 管理终端 I/F、212: RAID 控制控制器、213: 存储器、214: SAS 控制器、250: 宽端口、300: 控制程序、320: SAS 扩展器控制程序、330: RG/LU 设定程序、340: 区域构成控制程序、350: 性能监视程序、360: 后台控制程序、370: 故障处理程序

#### 具体实施方式

以下, 参照附图以应用了本发明的一种实施方式的计算机系统的存储系统为例进行说明。

图 1 是表示本发明的一种实施方式中的存储系统结构例的框图。以下, 对同种要素标记相同的主号码, 在对同种要素分别进行说明时, 在主号码上标记子符号来说明。

存储系统 100 能够通过 FC(Fiber Channel)、SCSI(Small Computer System

Interface)、SAS (Serial Attached SCSI)、IP (Internet Protocol) 等接口 103 与连接有一台或多台 (例如两台) 主机 101a、101b 的 SAN (存储区网络) 102 相连。另外, 存储系统 100 也可以与例如连接有管理终端 104 的管理用网络 106 相连。管理用网络 106 可以采用各种网络 (例如 LAN (Local Area Network))。也可以采用其他种类的网络代替 SAN102。另外, 管理用网络 106 和 SAN102 可以是一个通信网络。

各主机 101a、101b 是例如具备 CPU、存储器以及与存储系统 100 间的通信所使用的 I/O 端口等硬件资源的计算机。各主机 101a、101b 可以经由 SAN102 访问存储系统 100 内的数据。

管理终端 104 是具备 CPU、输入输出装置、存储器以及接口等硬件资源的计算机。管理终端 104 的 CPU 可以通过执行存储系统管理程序 105 来控制存储系统 100 (例如, 进行取得构成信息等操作指示)。

存储系统 100 具有: 多个可以存储数据的盘驱动器 114; 一个或多个 SAS 扩展器 130, 该 SAS 扩展器 130 是用于扩展连接盘驱动器 114 的交换机; 控制存储系统 100 的一个或多个存储控制控制器 200。盘驱动器 114 是硬盘驱动器, 但也可以采用 DVD (Digital Versatile Disk) 驱动器、CD (Compact Disk) 驱动器等其他种类的盘驱动器。在与 SAS 扩展器 130 相连的物理链接 108 中, 可以将使用不同协议的盘驱动器 114 与相同物理链接 108 相连。物理链接 108 可以转发驱动器 114 所支持的协议的帧。SAS 扩展器 130 可以由硬件电路板构成。

在图 1 中举例表示了三台盘驱动器 114a、114b、114c, 两个 SAS 扩展器 130a、130b 以及一个存储控制控制器 200。在以下的说明中适当采用此例进行说明。另外, 在以下的说明中将存储控制控制器 200 作为最上层。

两个 SAS 扩展器 130a、130b 被从上层至下层级联连接。在此例中, 由于 SAS 扩展器的数量为两个, 所以为串联连接, 但在 SAS 扩展器的数量为三个以上时, 也可以成为以 SAS 扩展器 103a 为起始的树形结构。

存储控制控制器 200 与位于其下方的 SAS 扩展器 103a 通过多个物理链接 108a、108b、108c 相连, 由这些物理链接 108a、108b、108c 形成了一个宽链接 120a。同样地, SAS 扩展器 130a 与 SAS 扩展器 130b 也通过多个物理

链接 108d、108e、108f 相连，由这些物理链接 108d、108e、108f 形成宽链接 120b。

为了识别连接物理链接的端口（以下称为物理端口），在各 SAS 扩展器 130a、130b 中分配了 ID。在本实施方式中，SAS 扩展器 130a 的物理端口号码 0 对应于物理链接 108a，SAS 扩展器 130a 的物理端口号码 1 对应于物理链接 108b，SAS 扩展器 130a 的物理端口号码 2 对应于物理链接 108c，SAS 扩展器 130a 的物理端口号码 3 对应于物理链接 108g，SAS 扩展器 130a 的物理端口号码 4 对应于物理链接 108d，SAS 扩展器 130a 的物理端口号码 5 对应于物理链接 108e，SAS 扩展器 130a 的物理端口号码 6 对应于物理链接 108f。同样地，SAS 扩展器 130b 物理端口号码 1 对应于物理链接 108d，SAS 扩展器 130b 物理端口号码 2 对应于物理链接 108e，SAS 扩展器 130b 物理端口号码 3 对应于物理链接 108f，SAS 扩展器 130b 物理端口号码 4 对应于物理链接 108h，SAS 扩展器 130b 物理端口号码 5 对应于物理链接 108i。

SAS 扩展器 130a 的物理端口号码 4~6 被设定为下层物理端口，SAS 扩展器 130b 的物理端口号码 1~3 被设定为上层物理端口。从 SAS 扩展器 130b 来看，SAS 扩展器 130a 是上层装置。

多个盘驱动器 114 可以仅由进行同种协议的通信（换言之，具有同种 I/F）的盘驱动器构成，也可以由不同协议的盘驱动器并存而构成。作为后者的例子，具有 SCSI I/F 的盘驱动器（以下称为 SCSI 驱动器）和具有 ATA I/F 的盘驱动器（以下称为 ATA 驱动器）可以并存。作为 SCSI 驱动器可以采用具有 SAS I/F 的盘驱动器。作为 ATA 驱动器可以采用具有 SATA I/F（或 PATA（Parallel ATA）I/F）的盘驱动器。此外，在采用了 PATA I/F 的驱动器时，在该驱动器与 SAS 扩展器 130 之间的通信中进行将串行传输变换为并行传输的处理（PATA 与 SATA 的物理层变换）。

出于存储系统 100 内的冗余化目的，盘驱动器 114 有时具有多个物理端口。例如，当盘驱动器 114 为 SATA 驱动器时，为了出于冗余化目的而增加端口数，有时使 SATA 端口选择器介于 SAS 扩展器 130 和 SATA 驱动器之间（在图 1 中未表示 SATA 端口选择器）。

盘驱动器 114a 经由物理链接 108g 与 SAS 扩展器 130a 的 3 号物理端口相

连。盘驱动器 114b 经由物理链接 108h 与 SAS 扩展器 130b 的 4 号物理端口相连。盘驱动器 114c 经由物理链接 108i 与 SAS 扩展器 130b 的 5 号物理端口相连。

盘驱动器 114b 和 114c 构成遵从 RAID 规则的组（以下称为 RAID 组）140。一个 RAID 组 140 可以由多个盘驱动器 114 构成，但这些盘驱动器 114 各自可以从属于一个存储控制控制器 200，与一个 SAS 扩展器 130 相连，也可以与不同 SAS 扩展器 130 相连，或者也可以从属于不同存储控制控制器 200。即，RAID 组 140 可以跨越多个 SAS 扩展器 130 而构成，也可以跨越多个存储控制控制器 200 而构成。

SAS 扩展器 130 具有存储区（例如存储器）。在该存储区内存储有区域设定程序 190、直接装置表（direct device table）195 以及扩展器路由表 196。在直接装置表 195 中记录与具有该表 195 的 SAS 扩展器 130 直接连接的装置的信息。另一方面，在扩展器路由表 196 中记录与具有该表 196 的 SAS 扩展器 130 间接连接的装置的信息。表 195、196 可以为一体。

区域设定程序 190 根据来自存储控制控制器 200 中 SAS 扩展器控制程序 320（图 3）的指示，变更直接装置表 195 和/或扩展器路由表 196 的内容，由此可以变更区域 150 的设定。关于直接装置表 195、扩展器路由表 196 以及区域 150 的详细情况，在后面进行叙述。

在存储系统 100 中可以设定一个或多个虚拟区域 150。存储系统 100 可以根据区域 150 的设定来控制物理链接的分配。区域 150 可以根据各种条件进行设定。具体而言，例如可以根据盘驱动器 114 的协议种类、或以 RAID 组 140 为单位来设定区域 150。

图 1 中举例表示的区域 150a 包含物理链接 108a 及 108b、物理链接 108d 及 108e、和物理链接 108h 及 108i。这种情况下，当存储控制控制器 200 访问 RAID 组 140 时，在宽链接 120a 中使用物理链接 108a 和 108b 中的某个，而不使用区域 150a 中不包含的物理链接 108c。

另一方面，图 1 中举例表示的区域 150b 包含物理链接 108c 及 108g。因此，当存储控制控制器 200 访问盘驱动器 114a 时，在宽链接 120a 中仅使用物理链接 108c，而不使用区域 150 中不包含的物理链接 108a 及 108b。

图 2 是表示存储控制控制器 200 的结构例的框图。

存储控制控制器 200 具备：主机 I/F 控制器 210、管理终端 I/F211、RAID 控制控制器 212、存储器 213 以及 SAS 控制器 214。主机 I/F 控制器 210、SAS 控制器 214、RAID 控制控制器 212 以及存储器 213 中的至少一个，不限于一个而可以为多个。

主机 I/F 控制器 210 通过图 1 的接口 103 经由 SAN102 与主机 101 可通信地连接。主机 I/F 控制器 210 可以控制主机 101 与 RAID 控制控制器 212 之间通信的协议变换。

管理终端 I/F211 经由图 1 的管理用网络 106 与管理终端 104 可通信地连接。管理终端 I/F211 可以控制管理终端 104 与 RAID 控制控制器 212 之间通信的协议变换。

主机 I/F 控制器 210、管理终端 I/F211、存储器 213 以及 SAS 控制器 214 分别通过总线 220a、220b、220c、220d 与 RAID 控制控制器 212 相连。总线 220 为例如 PCI 或 PCI-Express 等。

存储器 213 存储有控制程序 300，将其读入 RAID 控制控制器 212 并执行。关于控制程序 300 的详细情况在后面进行叙述。

SAS 控制器 214 可以控制 SAS 扩展器 130 与 RAID 控制控制器 212 之间通信的协议变换。SAS 控制器 214 具有宽端口 250。宽端口 250 是由两个以上物理端口构成的端口，经由宽链接 120a 与 SAS 扩展器 130a 相连。SAS 控制器 214 可以具有多个宽端口 250。换言之，当在 SAS 控制器 214 上连接多个 SAS 扩展器 130 时，SAS 控制器 214 具有多个宽端口 250。

RAID 控制控制器 212 可以具备例如：微处理器、进行数据转发的网桥（bridge）（例如 LSI（Large Scale Integration））以及 RAID 控制逻辑电路（例如异或 X-or 引擎（engine））等。RAID 控制控制器 212 内的处理器或网桥、RAID 控制逻辑电路等也可以划分在多个芯片上。

存储控制控制器 200，当从主机 101 接收写命令时，可以将遵从该写命令的数据暂时存储在存储器 213 中，将存储在存储器 213 中的数据写入盘驱动器 114。另一方面，存储控制控制器 200，当从主机 101 接收读命令时，可以从盘驱动器 114 读出遵从该读命令的数据，将读出的数据暂时存储在存储

器 213 中，将存储在存储器 213 中的数据发送至主机 101。

图 3 表示控制程序 300 的结构例。

控制程序 300 具有例如：SAS 扩展器控制程序 320、RG/LU 设定程序 330、区域构成控制程序 340、性能监视程序 350、后台（background）控制程序 360 以及故障处理程序 370。

SAS 扩展器控制程序 320 发出针对 SAS 扩展器 130 内的直接装置表 195 的查询，由此可以获取盘驱动器 114 以及 SAS 扩展器 130 的物理安装位置。

另外，SAS 扩展器控制程序 320 可以登录下层的间接连接装置的地址，在扩展器路由表 196 中设定路径信息。所谓间接连接装置，在 SAS 扩展器 130a 的扩展器路由表 196a 的情况下，是指可以不与该 SAS 扩展器 130a 的各物理端口直接连接，而从该物理端口通过至少一个其他物理端口进行访问的装置。具体而言，以图 1 的例子来说，0、1 号物理端口的间接连接装置是盘驱动器 114b 和 114c，2 号物理端口的间接连接装置是盘驱动器 114a。因此，例如为了确立与盘驱动器 114b 以及 114c 的连接，需要在 SAS 扩展器 130a 的扩展器路由表 196a 中登录用于到达盘驱动器 114b 以及 114c 的路径信息。

SAS 扩展器控制程序 320 可以获取构成宽链接 120 的多个物理链接 108 的数量、以及用于识别各物理链接的 ID（也可以是 SAS 扩展器的物理端口号码），并登录在直接装置表 195 中。换言之，SAS 扩展器 130 可以在 SAS 扩展器的直接装置表 195 中保存宽链接 120、构成宽链接 120 的物理链接 108 的数量、以及用于识别物理链接的 ID。

RG/LU 设定程序 330 是用于设定 RAID 组 140 且在 RAID 组 140 中设定 LU（逻辑单元）的程序。LU 也被称为逻辑卷。主机 101 发出指定了 LUN（逻辑单元号码）的 I/O 命令，该 LUN 是用于识别 LU 的 ID。存储控制控制器 200 可以使用由主机 101 指定的 LUN，来确定与该 LUN 对应的 RAID 组 140 并访问所确定的 RAID 组 140。

区域构成控制程序 340 是用于变更区域 150 的设定的程序。区域构成控制程序 340 可以决定将构成宽链接 120 的物理链接 108 逻辑地分配给盘驱动器 114 或 RAID 组 140 的设定，向 SAS 扩展器 130 的区域设定程序 190 请求区域变更。另外，区域构成控制程序 340，通过与由管理员操作的存储系统

管理程序 105 协同工作，也可以由管理员手动变更区域 150 的设定。

性能监视程序 350 是用于检测主机 101 或后台控制程序 360 的 I/O 请求、生成性能统计信息的程序。性能监视程序 350 可以根据过去的 I/O 特性的履历或检测到的性能来生成性能统计信息。

后台控制程序 360 是在后端执行的计算机程序。后台控制程序 360 是执行例如 RAID 组 140 的集合复制 (collection copy) 或、盘驱动器 114 间的数据转移或复制等操作的程序。后台控制程序 360 是与存储系统内部执行的 I/O 处理相关的程序的总称。以后，后台控制程序 360 所执行的 I/O，由于不是根据来自主机 101 的 I/O 请求而进行的 I/O，因此被称为内部 I/O。另外，数据的复制或转移不限于存储系统 100 内的复制或转移，可以是例如向外部存储系统内的盘驱动器的复制 (所谓的远程复制) 或转移。

故障处理程序 370 是检测存储系统 100 内各种故障的程序。具体而言，例如，故障处理程序 370 构成为，当构成宽链接 120 的物理链接 108 中发生了故障时可以检测出该故障并执行对应于该故障的处理。

接下来说明直接装置表 195 和扩展器路由表 196，此时对区域的控制进行说明。

图 4 表示 SAS 扩展器 130a 内的直接装置表 195a 的结构例以及在该表 195a 中登录的信息。图 5 表示 SAS 扩展器 130b 内的直接装置表 195b 的结构例以及在该表 195b 中登录的信息。图 6 表示 SAS 扩展器 130a 内的扩展器路由表 196a 的结构例以及在该表 196a 中登录的信息。图 7 表示 SAS 扩展器 130b 的扩展器路由表 196b 的结构例以及在该表 196b 中登录的信息。

直接装置表 195 是用于将与 SAS 扩展器 130 直接相连的装置 (驱动器、SAS 控制器等) 与 SAS 扩展器 130 的物理端口号码相关联的表。另外，直接装置表 195 中也保存用于管理区域设定的信息。

对图 4 进行说明。在直接装置表 195a 的字段 (field) 410 中，记录着与 SAS 扩展器 130a 的多个物理端口分别对应的多个物理端口号码。在字段 420 中，针对 SAS 扩展器 130a 的每个物理端口，记录着遵从与该物理端口直接相连的装置的 SAS 的地址 (以下称为 SAS 地址)。在字段 430 中，针对 SAS 扩展器 130a 的每个物理端口，记录着许可通过的装置 (以下称为许可装置)

的 SAS 地址。许可装置是存储系统 100 的后端的访问源。访问源根据从上层至下层发送数据还是从下层至上层发送数据而不同，在前一种情况下访问源为 SAS 控制器 214，在后一种情况下访问源为盘驱动器 114。所以在本实施方式中，许可装置为 SAS 控制器 214 或盘驱动器 114 中的某一种。

在图 1 中举例表示的连接结构的情况下，与 SAS 扩展器 130a 的物理端口号码 0 的物理端口直接相连的装置是 SAS 控制器 214。从而，SAS 扩展器 130a 在字段 420 中对应于物理端口号码 0 的栏中，登录 SAS 控制器 214 的 SAS 地址。SAS 扩展器 130 对于其他物理端口号码也同样地登录图 1 中举例表示的连接结构的信息。

另外，在图 1 的区域 150a 的情况下，使用 SAS 扩展器 130a 的物理端口号码 0 的装置为盘驱动器 114b 以及 114c 中的某一个。从而，区域构成控制程序 340，作为对应于物理端口号码 0 的许可装置，在字段 430 中登录盘驱动器 114b 以及 114c 的 SAS 地址。区域构成控制程序 340，对于其他物理端口号码也同样地如图 1 的区域 150a 以及 150b 那样，登录对应于各物理端口号码的许可装置的 SAS 地址。此外，字段 430 中的 NULL 表示没有许可装置。从而，如 SAS 扩展器 130a 的物理端口号码 6 那样，许可装置全部不存在的情况下（参照图 1），在许可装置栏（字段 430 中的栏）中全部记录 NULL。

另外，在直接装置表 195a 中可以将宽链接与构成该宽链接的多个物理链接关联起来。此时，可以设定表示该宽链接对于具有该表 195a 的 SAS 扩展器 130a 而言是否在上层的信息。由此，SAS 扩展器 130a 可以判断，经过哪个宽链接使来自盘驱动器 140 的数据向更上层流动较好。根据该表 195a，当从上层的物理端口号码 2 的物理端口接收到向盘驱动器 114a 的 I/O 时，SAS 扩展器 130a 根据表 195a 得知盘驱动器 114a 与物理端口号码 3 的物理端口直接相连，因此可以从物理端口号码 3 的物理端口开始进行 I/O。

对图 5 进行说明（此时，对与图 4 说明的共同点省略或简化说明。）

在图 5 的表 195b 中也登录有表示图 1 举例所示的连接结构的信息。即，例如与 SAS 扩展器 130b 的物理端口号码 0 的物理端口直接连接的是 SAS 扩展器 130a，因此在与该物理端口号码 0 相对应的字段 410 的栏中登录 SAS 扩展器 130a 的 SAS 地址。另外，例如，由于不存在与 SAS 扩展器 130b 的物理



端口号码 5、6 的物理端口直接连接的装置，所以在与该物理端口号码 5、6 分别对应的字段 410 的栏中登录 NULL。根据该表 195b，当从上层的物理端口号码 1 的物理端口接收到向盘驱动器 114b 的 I/O 时，SAS 扩展器 130b 根据表 195b 得知盘驱动器 114b 与物理端口号码 3 的物理端口直接连接，因此可以从物理端口号码 3 的物理端口开始进行 I/O。

如上所述，在本实施方式中，在直接装置表 195 中针对具有该表 195 的 SAS 扩展器 130 的每个物理端口记录有：物理端口号码（也可以是其他种类的 ID）、表示与该物理端口直接连接的装置的 SAS 地址（也可以是其他种类的信息）、表示被许可经过该物理端口的访问源装置的 SAS 地址（也可以是其他种类的信息）。

接着对扩展器路由表 196 进行说明。扩展器路由表 196 保存与具有该表 196 的 SAS 扩展器 130 的下层间接连接（至少间隔一个物理端口来连接）的装置的信息。以下将指定了 SAS 标准中 Subtractive 属性的 SAS 扩展器 130 的宽端口 250（Logical port）称为“下层宽端口”，另外，以下将指定了 SAS 标准中 Table 属性的 SAS 扩展器 130 的宽端口称为“上层宽端口”。关于扩展器路由表 196 的登录内容，按照 SAS 标准的扩展器路由表，在某个 SAS 扩展器 130 的下层宽端口上连接其他 SAS 扩展器 130，当间接连接装置与所述其他 SAS 扩展器 130 相连时，SAS 扩展器控制程序 320 需要在上层的 SAS 扩展器 130 的扩展器路由表 196 中登录间接连接装置的 SAS 地址。

对图 6 进行说明。在图 1 中，SAS 扩展器 130a 在其下层宽端口上连接有 SAS 扩展器 130b，因此 SAS 扩展器控制程序 320 在扩展器路由表 196a 的、与构成宽链接 120a 的各物理端口的号码相对应的字段 510 中，记录经由图 1 的宽链接 120a 访问的间接连接装置（盘驱动器 114b、114c）的 SAS 地址（在图 6 中，“Physical Phy”表示物理端口）。根据该表 196a，当从上层接收到对盘驱动器 114b 的 I/O 时，SAS 扩展器 130a 根据表 195a 以及 196a 得知盘驱动器 114b 为间接连接装置，并将其分配给下层的物理端口号码 4 或 5，因此可以从物理端口号码 4 或 5 的物理端口开始进行 I/O。

对图 7 进行说明。在图 1 中，SAS 扩展器 130b 不在其下层宽端口上连接其他 SAS 扩展器 130。因此，扩展器路由表 196b 的字段 510 为全部未登录。

通过以上在图 4~图 7 中举例所示的各表结构,例如可以控制从存储控制控制器 200 向各盘驱动器 114 的访问。以下参照图 21 并适当地进一步参照图 1、图 2 以及图 4~图 7 来分别进行说明。

存储控制控制器 200,当访问希望的盘驱动器 114 时发送用于确立连接的规定帧(以下称为打开地址帧(Open Address Frame)),然后,当接收到表示连接确立的应答时可以访问该希望的盘驱动器 114。

具体而言,例如存储控制控制器 200 在盘驱动器(以下在此为 SAS 目标装置) 114b 中写入数据时,首先将控制程序 300 发送帧(包含写入请求以及写入对象数据的帧)的请求发送至 SAS 控制器(以下,在此为 SAS 启动器装置) 214 (S0)。SAS 启动器装置 214 接收该请求,通过从宽端口 250 的任意物理端口号码的物理端口发送针对 SAS 目标装置 114b 的打开地址帧,尝试确立连接(S1)。在此阶段,SAS 启动器装置 214 按照规定规则指定物理端口号码,由于 SAS 协议的制约而无法以不遵从该规则的方法选择物理端口号码。例如,若该规则为按照小号码顺序来指定物理端口号码,则 SAS 启动器装置 214 首先从物理端口号码 0 的物理端口发送打开地址帧,不可以无视该规则地从第一次起例如在物理端口 0、1 未确立连接的状态下从物理端口号码 2 的物理端口发送打开地址帧。

当接收到打开地址帧时,SAS 扩展器 130a 参照扩展器路由表 196a,查找用于与 SAS 目标装置 114b 确立连接的路径(S2)。在此,SAS 扩展器 130a,当通过宽端口配备有多个物理端口时,除了由于其他连接或区域设定而无法使用的物理端口以外,选择空闲物理端口中的小号码。

SAS 扩展器 130a,当参照直接装置表 195a 或扩展器路由表 196a 而得到的结果为未找到路径、没有目标、因区域未登录而不可连接中的任意一种情况时,向 SAS 启动器装置 214 返回 OPEN REJECT 错误。当接收到 OPEN REJECT 错误时(S4),SAS 启动器装置 214 将该错误中所包含的信息例如 SAS 目标装置的地址错误、或区域未分配而导致的不可确立连接发送至控制程序 300,控制程序 300 对其进行接收(S6)。当区域未分配而导致连接确立失败时,在 SAS 的被称为 Primitive 的信号中新定义表示因区域未分配而导致的连接确立失败的 Primitive,SAS 扩展器发送新定义的所述 Primitive,由此

也可以识别因区域未分配而导致的不可确立连接。

当 SAS 扩展器 130a 通过物理端口号码 2 的物理端口接收到打开地址帧时，进行该段落中的一连串的处理。这是由于在直接装置表 195a 中，在对应于物理端口号码 2 的栏中未登录盘驱动器 114b 的 SAS 地址，并且在扩展器路由表 196a 中，在对应于物理端口号码 2 的栏中也未登录盘驱动器 114b 的 SAS 地址。

在 S2 中，当不是未找到路径、没有目标、因区域未登录而不可连接中的任意一种情况，但构成宽端口的全部物理端口正处于其他连接的使用过程中时，SAS 扩展器 130a 向 SAS 启动器装置 214 返回 OPEN REJECT。SAS 启动器装置 214 当接收到 OPNE REJECT 时 (S5) 再次执行 S1。重试 (retry) 次数以及重试所需时间由 SAS 启动器装置 214 管理，当达到阈值时向控制程序通知重试失败。

在 S2 中，不是未找到路径、没有目标、因未登录区域而不可连接的任意一种情况，且存在遵从所找到的路径的空闲物理端口时，SAS 扩展器 130a 选择该空闲物理端口，确定地址帧输入侧的 SAS 扩展器的物理端口和所述空闲物理端口的路径（确定 SAS 扩展器的内部切换），通过所选择的空闲物理端口，向下层的 SAS 扩展器 130b 转发地址帧 (S3)。由此，在 SAS 扩展器 130b 中也进行与 SAS 扩展器 130a 相同的 S2 处理，当找到 SAS 目标装置 114b 时，在 SAS 扩展器 130b 与 SAS 目标装置 114b 之间进行通信，确立连接 (S7)。SAS 目标装置 114b 发出 OPEN Accept, OPEN Accept 通过在 SAS 启动器装置 214 与 SAS 目标装置 114b 之间所确立的连接（路径）到达 SAS 启动器装置 214 (S8)。SAS 启动器装置 214 对 OPEN Accept 的接收进行应答，向 SAS 目标装置 114b 转发遵从在 S1 中接收到的帧发送请求的帧 (S9)。SAS 目标装置 114b 当接收到帧时 (S10) 返回 ACK。SAS 启动器装置 214 接收 ACK (S11)，当不需要转发帧时执行连接结束步骤 (S12)，通过释放全部 SAS 扩展器的物理端口来结束连接。

以上是确立连接时进行的处理流程的一例。当从 SAS 目标装置向 SAS 启动器装置发送数据时，上述 SAS 目标装置对 SAS 启动器装置发出连接请求，从 SAS 目标装置向 SAS 启动器装置转发帧。除了由 SAS 目标装置执行连接

请求这一点以外与图 21 的步骤图相同，因此省略说明。

以下说明本实施方式的存储系统 100 中进行的处理流程的一例。

图 8 表示存储系统启动时所进行的处理流程的一例。此外，图中的“RG”表示 RAID 组。

RG/LU 设定程序 330 取得已设定好的 RAID 组 140 的信息（步骤 400）。例如可以从 RAID 组管理表 509 取得该信息。RAID 组管理表 509 中，针对每个 RAID 组记录有表示该 RAID 组的号码、构成该 RAID 组的盘驱动器 114 的协议种类。虽未图示，但除此以外也可以针对每个 RAID 组记录表示构成 RAID 组的盘驱动器的号码、或在 RAID 组上所设定的 LU 的 LUN 等其他种类的信息。

区域构成控制程序 340 将对应于全部 RAID 组 140 的区域 150 初始化（例如，在每个 RAID 组中设定一个区域 150）（步骤 410），将物理链接分配给各区域 150（步骤 415）。

此外，在此启动时，当预先在管理终端 104 中存在根据协议的种类来分配所使用的物理链接的指示时，在步骤 415 中，区域构成控制程序 340 根据协议的种类，不重叠地（overlap）划分区域 150。在图 1 的例子中，区域 150a 包含构成宽链接 120a 的物理链接 108a、108b，区域 150b 包含构成宽链接 120a 的物理链接 108c，在双方区域中没有共用的物理链接 108 地设定了区域 150。可以使用构成 RAID 组 140 的驱动器种类、驱动器的性能、驱动器的台数、以性能监视程序 350 监视的 RAID 组 140 的 I/O 请求性能的统计信息等各种条件来计算构成宽链接 120 的物理链接数。

图 9 表示对分配给区域 150 的物理链接 108 的数量进行变更的处理流程的一例。

根据来自管理终端 104 的指示，开始执行性能监视程序 350。当性能监视的指示到来时，性能监视程序 350 开始执行针对各 RAID 组 140 的性能统计信息的获取并前进至步骤 520。当指示未到来时，保持在步骤 510 中等待指示。此外，在此所说的“性能”是每单位时间转发的数据大小。

在步骤 520 中，若从管理终端 104 发来性能监视结束的指示，则性能监视程序 350 结束执行，并结束流程。在没有结束指示的情况下，前进至步骤

530。

在步骤 530 中，根据性能统计信息，当超过了监视对象 RAID 组 140 的性能阈值时前进至步骤 600，若步骤 600 结束则返回步骤 520。在步骤 530 中使用的性能阈值，例如是根据区域 150 包含的物理链接 108 的数量而确定的值。

图 10 表示图 9 的步骤 600 中进行的处理流程的一例。

在步骤 610 中，性能监视程序 350 根据性能统计信息计算出需要的物理链接 108 的数量。具体而言，例如，计算出适合于超过性能阈值的性能的数量、或者适合于根据性能统计信息推测出的今后的最大性能的数量。

在步骤 630 中，区域构成控制程序 340 判断能否在变更的区域 150 中再追加物理链接 108。当能够再追加物理链接时前进到步骤 640，当不能追加时前进到步骤 650。在此，所谓可以追加的情况是指例如存在未分配给任何区域的物理链接（或者已经分配给区域，也可以再分配给其他区域的物理链接）的情况，所谓不能追加的情况是指不存在未分配给区域的物理链接（或者已经分配给区域，不分配给其他区域的物理链接）的情况。

在步骤 640 中，区域构成控制程序 340，当变更的区域与其他区域 150 共享宽链接 120 时（例如图 1 所示，两个区域 150a、150b 共享一个宽链接 120a 时），考虑与其他 RAID 组 140（或盘驱动器 114）相关的性能信息，计算出可以不对其他 RAID 组 140（或盘驱动器 114）的性能造成不良影响地向变更对象区域 150 的宽链接 120 追加的物理链接 108 的数量。

在步骤 650 中，区域构成控制程序 340 判断与想分配给物理链接 108 的区域 150 不同的区域 150 是否共享着物理链接 108。在共享着的情况下，区域构成控制程序 340 检查其他区域 150 的物理链接 108 的数量是否为最小值（仅含有一个物理链接 108）。当不是最小值时，由于存在减少其他区域 150 的物理链接 108 的数量的余地，因此区域构成控制程序 340 前进到步骤 660。当是最小值时，由于所有区域 150 均无法变更，因此前进到步骤 680。

在步骤 660 中，区域构成控制程序 340 关于其他区域 150 计算削减的物理链接数，前进到步骤 670。

在步骤 670 中，区域构成控制程序 340 变更区域 150 中的物理链接 108

的数量。

关于该步骤 670，以在区域 150 中增加物理链接的情况为例加以说明。例如，区域构成控制程序 340，在图 1 中当未对 RAID 组 140（盘驱动器 114b、114c 的集合）分配物理链接 108a 时，若将物理链接 108a 分配给 RAID 组 140，则对 SAS 扩展器 130a 发出物理链接追加指示。在该物理链接追加指示中包含有 SAS 扩展器 130a 的物理端口号码 0、分配对象装置的地址（盘驱动器 114b、114c 各自的 SAS 地址）。区域设定程序 190a，根据该物理链接追加指示，在对应于直接装置表 195a 的物理端口号码 0 的字段 430 栏中记录分配对象装置的地址（盘驱动器 114b、114c 各自的 SAS 地址）。并且，区域设定程序 190a 在对应于扩展器路由表 196a 的物理端口号码 0 的栏中，记录盘驱动器 114b、114c 各自的 SAS 地址。

关于步骤 670，以从区域 150 删除已分配的物理链接的情况为例加以说明。例如，在图 1 中，区域构成控制程序 340 在对 RAID 组 140（盘驱动器 114b、114c 的集合）已分配了物理链接 108a 的情况下，若解除该物理链接 108a 的分配，则对 SAS 扩展器 130a 发出物理链接删除指示。在该物理链接删除指示中包含 SAS 扩展器 130a 的物理端口号码 0、分配解除对象装置的地址（盘驱动器 114b、114c 各自的 SAS 地址）。区域设定程序 190a 根据该物理链接删除指示，从对应于直接装置表 195a 的物理端口号码 0 的字段 430 栏中删除盘驱动器 114b、114c 各自的 SAS 地址，并记录 NULL。并且，区域设定程序 190a 从对应于扩展器路由表 196a 的物理端口号码 0 的栏中删除分配对象装置的地址（盘驱动器 114b、114c 各自的 SAS 地址），成为未登录。

如上所述，分配物理链接或解除该分配的处理，针对 SAS 扩展器 130 中的直接装置表 195 以及扩展器路由表 196 中的至少一个，追加成为物理链接分配对象的装置（SAS 目标）的信息、或者删除成为物理链接分配解除对象的装置的信息，所述 SAS 扩展器 130 具有作为分配对象或分配解除对象的物理链接。控制程序 300 掌握存在于存储控制控制器 200 的下层的 SAS 扩展器 130、或各 SAS 扩展器 130 的物理端口号码（例如，保存全部 SAS 扩展器 130 的 SAS 地址或各 SAS 扩展器 130 的物理端口号码）。控制程序 300 可以对具有分配对象或分配解除对象的物理链接的 SAS 扩展器 130 指定该物理链接所

连接的物理端口的物理端口号码、或与该物理端口号码新对应的或从该物理端口号码删除的 SAS 地址等。在删除全部 SAS 地址的情况下，也可以代替 SAS 地址而指定该意思。

在步骤 680 中，区域构成控制程序 340，由于无法对判断出性能为必要的区域 150 以及与该区域 150 共享的任何区域 150 变更物理链接 108，因此向管理终端 104 通知表示无法变更区域 150 的信息。

图 11 表示区域状态管理表 700 的结构例。在该图中，“RG”表示 RAID 组，“WL”表示宽链接，“phy”表示物理端口。

该区域状态管理表 700 是例如控制程序 300 所保存的表，包含例如以下信息。对各信息进行说明。

RG 号码 710 是 RAID 组 140 的识别号码。驱动器种类 720 是构成 RAID 组 140 的驱动器的种类（例如可以包含协议的种类、是否是存在性能差异的驱动器等其他类型的信息）。LU 号码 730 是存储对应于 RG 号码 710 的数据的逻辑单元的号码列表。区域号码 740 是对应于 RAID 组 140 的区域 150 的识别号码。性能的统计信息 750 保存由性能监视程序 350 取得的性能的统计信息。性能的阈值 760 是根据重叠的区域 150 的信息 790 计算出的值（阈值）。当从主机发送来超过该值的 I/O 请求时，在图 9 的步骤 530 中超过阈值，执行变更区域 150 的流程 600。构成 WL 的物理端口号码 770 存储有构成宽链接 120 的 SAS 扩展器 130 的物理端口号码。宽链接发生局部故障 780 是当构成宽链接 120 的物理链接 108 的一部分发生了故障时，登录在哪个物理链接 108 中发生了故障的列表。在正常状态时或从故障恢复后登录“未发生”。

控制程序 300，例如可以应答来自管理终端 104 的存储系统管理程序 105 的请求，将该表 700 提供给管理终端 104。存储系统管理程序 105 可以在管理终端 104 的显示装置上显示该表 700 所表示的信息。

以下说明本实施方式中进行物理链接的分配变更的情况的若干个例子。此外，由上述说明可知可以将构成一个宽链接的多个物理链接分配给多个区域，但以下将一个区域中一个宽链接内的一个以上物理链接称为“子宽链接”。

图 12A 表示进行物理链接的分配变更的第一情况中的分配变更前的物理链接的分配。图 12B 表示该分配变更后的物理链接的分配。

在物理链接的分配变更前，如图 12A 中举例所示，对一个 RAID 组 140 分配一个区域 150（例如，对 140a 分配 150a，对 140b 分配 150b）。另外，构成 RAID 组 140a 和 140b 的驱动器的种类不同（例如，140a 是 SAS 驱动器的集合，140b 是 SATA 驱动器的集合）。一个宽链接 120 由 4 个物理链接构成。并且，区域 150a 的子宽链接内的物理链接，和区域 150b 的子宽链接内的物理链接都是两个。

在这种情况下，性能监视程序 350 将来自主机 101a 的 I/O 请求与来自主机 101b 的 I/O 请求进行比较，判断为需要保持（sustain）性能。此时，性能监视程序 350 判断为针对区域 150a 需要将构成宽链接 120 的物理链接 108 的数量设定得较多。

因此，如图 12B 所示，区域构成控制程序 340 将区域 150b 的子宽链接内的物理链接减少一个，将这部分追加到区域 150a 的子宽链接上。这里，区域构成控制程序 340 进行物理链接的分配变更以使区域 150a 与区域 150b 不重叠。这是由于考虑到 RAID 组 140a 与 RAID 组 140b 的驱动器种类的不同，判断为使区域 150a 与区域 150b 不重叠。

由此，控制程序 300 可以将构成宽链接 120 的多个物理链接分配给驱动器或 RAID 组。分配物理链接的数量，通过监视来自主机的 I/O 请求的请求性能（即负荷）的增大，进行物理链接的分配控制，由此可以实现存储系统内的频带保证。另外，可以缓和由于驱动器的种类的不同而产生的对吞吐性能的影响。

图 13 表示进行物理链接的分配变更的第二情况中的分配变更前的物理链接的分配。图 14 表示该分配变更后的物理链接的分配。

在物理链接的分配变更前，如图 13 举例所示，对一个 RAID 组 140 分配一个区域 150。一个宽链接 120 由 5 个物理链接构成。并且区域 150a 的子宽链接内的物理链接为一个，区域 150b 的子宽链接内的物理链接为两个，区域 150c 的子宽链接内的物理链接为两个。通过后台控制程序 360 进行从 RAID 组 140b 向 RAID 组 140c 的数据复制（即，进行从 RAID 组 140b 读取数据、并将读出的数据写入 RAID 组 140c 的处理）。

在这种情况下，如图 14 所示，性能监视程序 350 判断为有来自主机 101



的对 RAID 组 140a 的 I/O 请求(或该 I/O 请求中的请求性能超过了规定阈值)。

因此,如图 14 所示,区域构成控制程序 340 将区域 150b 的子宽链接内的物理链接减少一个,并且将区域 150c 的子宽链接内的物理链接也减少一个,将这两部分追加到区域 150a 的子宽链接中。

由此,在存储系统内可以实现用于尽量满足来自主机的 I/O 请求中的性能的带宽保证。此外,虽然从区域 150b 以及 150c 的双方的子宽链接中减少了物理链接数,但如果内部 I/O 中也请求了与来自主机 101 的 I/O 请求相同程度的性能,也可以不从双方的子宽链接中减少物理连接数。减少或追加的物理链接数,如上所述可以是对应于各个 I/O 中所请求的性能或驱动器种类的适当值。

图 15A 表示进行物理链接的分配变更的第三情况中的分配变更前的物理链接的分配,图 15B 表示该分配变更后的物理链接的分配。

在物理链接的分配变更前,如图 15A 举例所示,对一个 RAID 组 140 分配一个区域 150。一个宽链接 120 由 4 个物理链接构成。并且区域 150a 的子宽链接内的物理链接为 3 个,区域 150b 的子宽链接内的物理链接为 1 个。由此可以处理来自主机 101a 的针对 RAID 组 140a 的 I/O 请求和来自主机 101b 的针对 RAID 组 140b 的 I/O 请求。

在这种情况下,故障处理程序 370 检测出在构成区域 150b 的子宽链接的物理链接中发生了故障。在这种状态下,无法从存储控制控制器 200 对 RAID 组 140b 进行访问,所以也无法处理来自主机 101b 的 I/O 请求。

因此,区域构成控制程序 340 使区域 150b 的子宽链接恢复到与故障发生前相同的状态地进行物理链接的分配变更。具体而言,如图 15B 所示,区域构成控制程序 340 将区域 150a 的子宽链接内的物理链接减少 1 个,将该部分追加到区域 150b 的子宽链接中。

由此,即使在构成宽链接的物理链接中发生故障,通过修改物理链接数的分配,也可以维持向盘驱动器或 RAID 组的访问路径。

在上述的实施方式中也可以进行以下操作。

如图 16 举例所示,可以使多个 RAID 组 140a、140b 与一个区域 150a 对应。在这种情况下,存储控制控制器 200 可以在区域 150a 的子宽链接的范围

内，对宽链接 120 中针对 RAID 组 140a 的物理链接数、和针对 RAID 组 140b 的物理链接数进行分配变更，不将其他区域 150b 的子宽链接内的物理链接追加至区域 150a 的子宽链接。另外，在区域 150a 的子宽链接的范围内进行分配变更时，在不对属于该区域 150a 的其他 RAID 组 140b 造成不良影响的范围内，可以调整分配给 RAID 组 140a 的物理链接数。可以将满足针对 RAID 组 140b 的请求性能的数量作为不造成不良影响的范围下限。

另外，尽管前面进行了叙述，但如图 17 所示，多个区域 150a、150b 可以在宽链接 120 中重复。即根据图 17 的例子，存储控制控制器 200 从宽链接 120 可以访问 RAID 组 140a 和另一 RAID 组 140b。

另外，如图 18 举例所示，可以没有所谓区域的概念。根据图 18 的例子，对于将宽链接 120 中的多个物理链接中的哪个物理链接分配给哪个 RAID 组 140（或盘驱动器）进行控制。

另外如图 19 举例所示，在存储系统 100 的外部访问源 1000a、1000b 中设定了优先度，可以根据访问目的地 RAID 组（或盘驱动器）的外部访问源的优先度的高低来决定分配给 RAID 组（或盘驱动器）的物理链接的数量。具体而言，例如当存在“大”和“小”两种优先度时，存储控制控制器 200 在宽链接 120 中，相对于来自优先度小的外部访问源 1000b 的访问目的地 RAID 组 140b，可以对来自优先度大的外部访问源 1000a 的访问目的地 RAID 组 140a 分配两倍数量的物理链接（例如可以对 140a 分配 4 个物理链接，对 140b 分配两个物理链接）。此外，在该变形例中，所谓外部访问源是主机 101 或该主机 101 内的应用程序（在操作系统上运行的计算机程序）。存储控制控制器 200 针对每个外部访问源，可以将表示该外部访问源的优先度的信息预先存储在存储器中。存储控制控制器 200，例如当确立与外部访问源的连接时或接收到 I/O 请求时，确定外部访问源并从存储器中检测对应于所确定的外部访问源的优先度，由此可以如上所述变更分配给所确定的外部访问源的访问目的地的物理链接数。

如图 20 举例所示，存储控制控制器 200 在宽链接 120 中分配给 RAID 组 140b（或盘驱动器）的物理链接数可以为零。这样一来，从存储控制控制器 200 无法访问 RAID 组 140b（或盘驱动器），但由此可以更新 RAID 组 140b

内的数据，可以不从 RAID 组 140a 读取数据。即当 RAID 组 140b 中要求高安全等级时，通过如上述那样使物理链接数为零，可以保护 RAID 组 140b 内的数据。此外，代替该方法，可以解除将 RAID 组 140b 与 SAS 扩展器 130 相连的物理链接的分配，但如上所述，较为理想的是使宽链接 120 中的物理链接为零，因为可分配给其他 RAID 组（或盘驱动器）的物理链接数可以增加该部分。另外，在上述说明中提到了 RAID 组的安全等级，但也可以是构成 RAID 组的盘驱动器或 LU 的安全等级。

以上说明了本发明的适当的实施方式以及若干变形例，但它们是用于说明本发明的例子，本发明的范围不限于这些实施方式以及变形例。也可以通过其他各种形态实施本发明。

例如，上述的存储系统也可以应用于服务器系统。此时，例如可以将存储控制控制器 200 作为服务器计算机（例如所谓的刀片服务器（blade server））。

另外，多个盘驱动器可以安装在机架上，性能监视程序 350 将用于定期监视该机架的资源（例如盘驱动器）的资源监视命令定期发送至规定目的地，由此可以定期监视其资源。此时可以设有该资源监视命令专用的区域。由此，可以将处理针对盘驱动器的 I/O 请求时所使用的物理链接、与资源监视命令的发送所使用的物理链接分离，从而可以尽量不给来自主机的 I/O 请求造成影响地进行资源监视。

另外，本发明的实施方式的计算机系统（例如存储系统）以使用 SAS I/F 为前提，但本发明的实施方式也可以应用于具有相当于 SAS 宽链接的功能的未来的 I/F。

另外，例如在所述说明中，被分配控制的“物理链接”可以是“逻辑链接”。也可以和上述的物理链接相同地，通过直接装置表和扩展器路由表对逻辑链接进行管理。所谓逻辑链接是通过对一个物理链接分时而形成的链接。具体而言，例如，慢转发率（例如 1.5Gbps）和快转发率（例如 3Gbps）并存的连接中的转发情况下，可以对速度快的物理链接（例如 3Gbps）进行分时，使某时间成为转发状态，而使本来可以转发的其他时间成为非转发状态。更具体地讲，如图 22 举例所示，通过将物理链接分时而为 2 个或 4 个，将转发率慢的多个逻辑链接聚集在一个物理链接中，在以低速连接有多台盘驱动器的

---

环境中，通过同时确立多个低速连接可以获得系统整体的转发速度。针对各逻辑链接，通过将表示直接装置表以及扩展器路由表的物理端口（phy）的索引（index）制成逻辑 phy 的表，可以进行所述的区域控制（在各区域中对逻辑链接进行分配控制）。

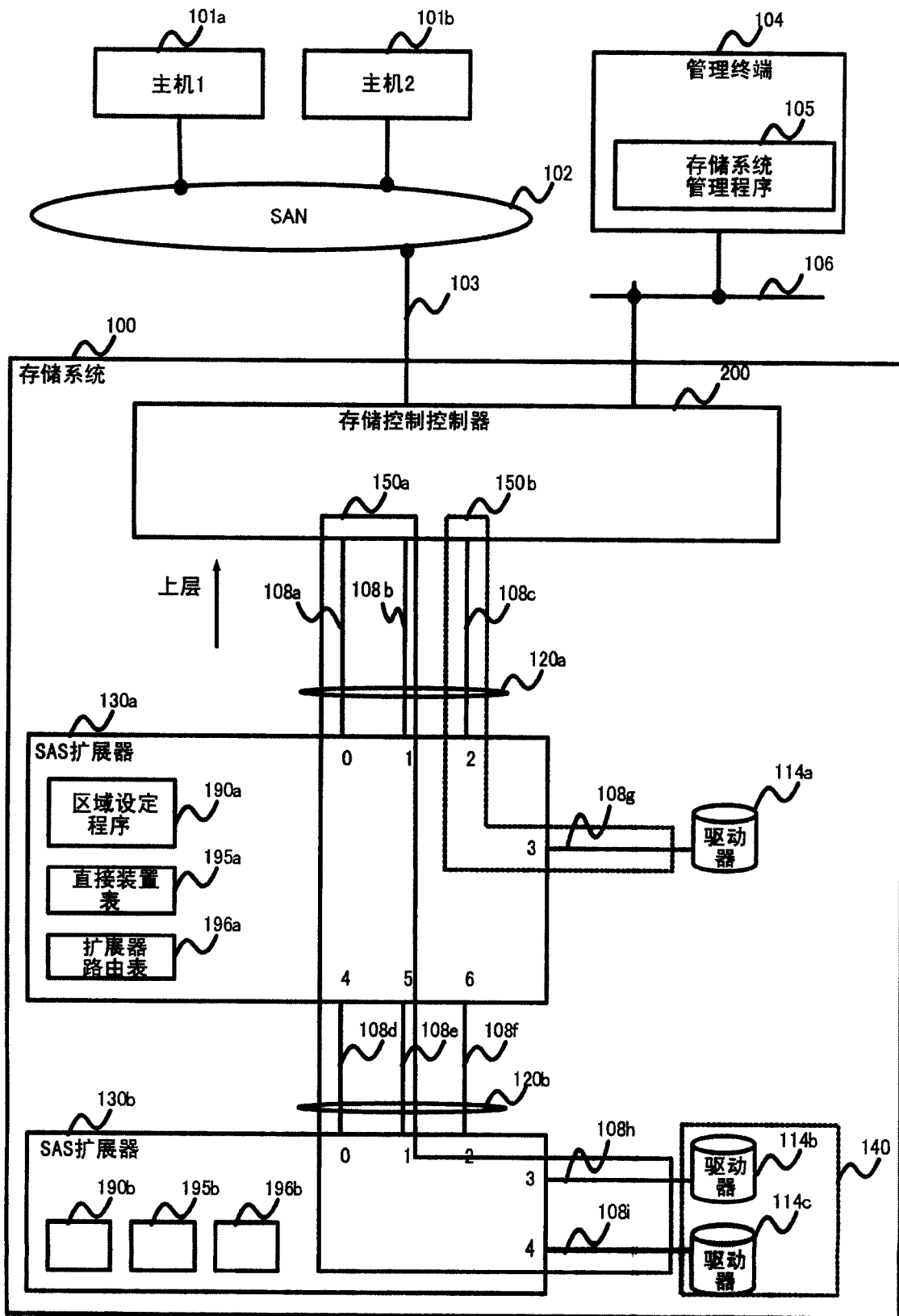


图 1

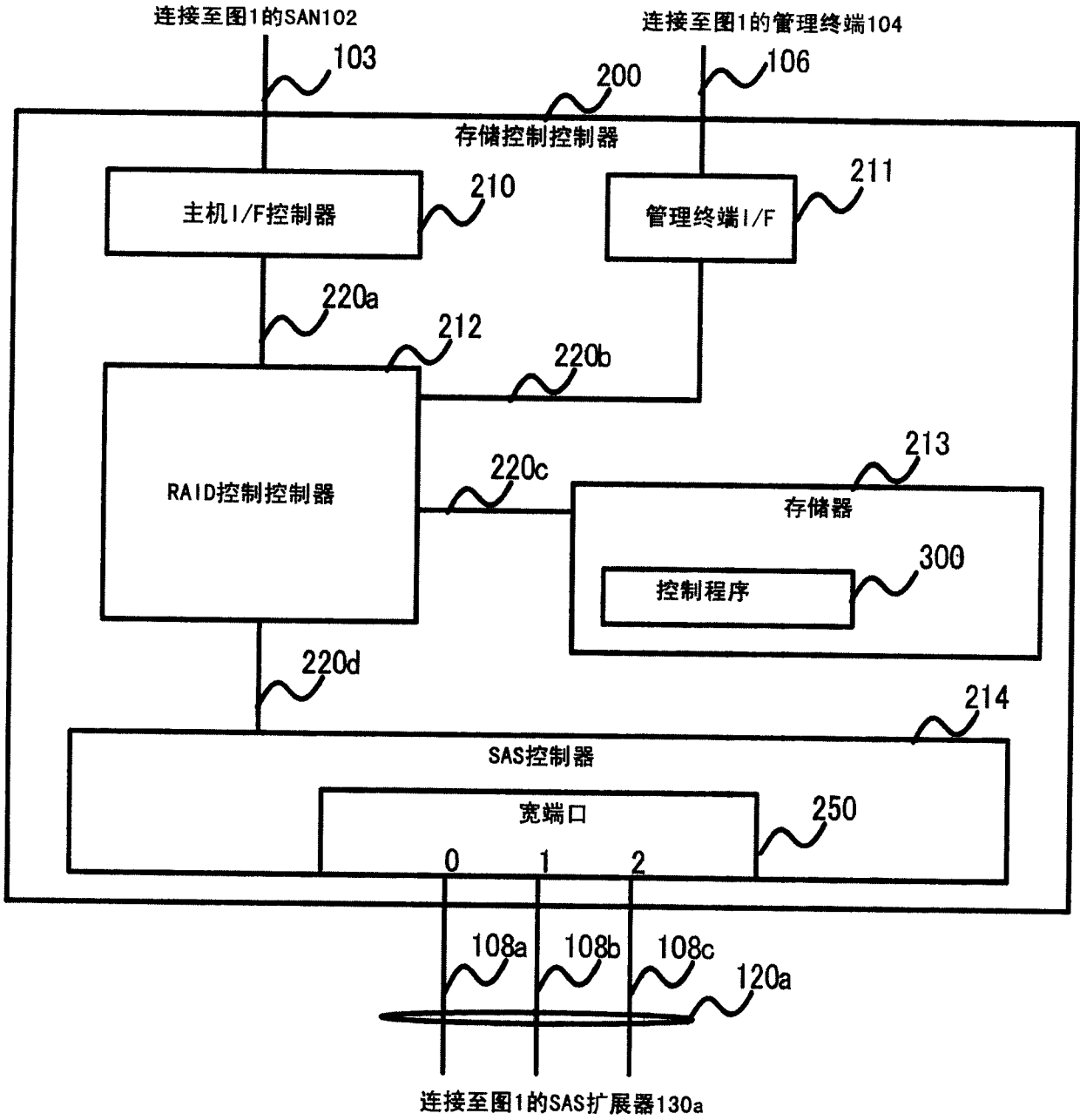


图 2

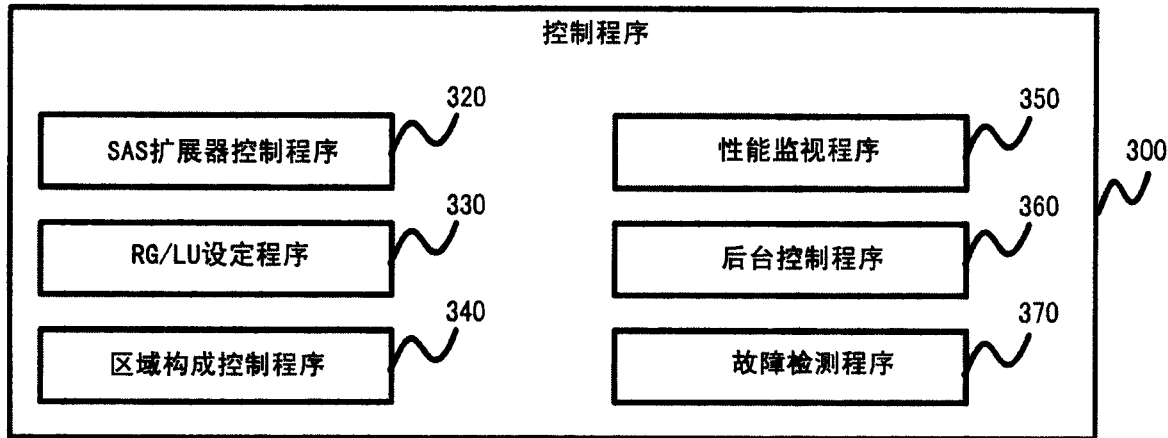


图 3

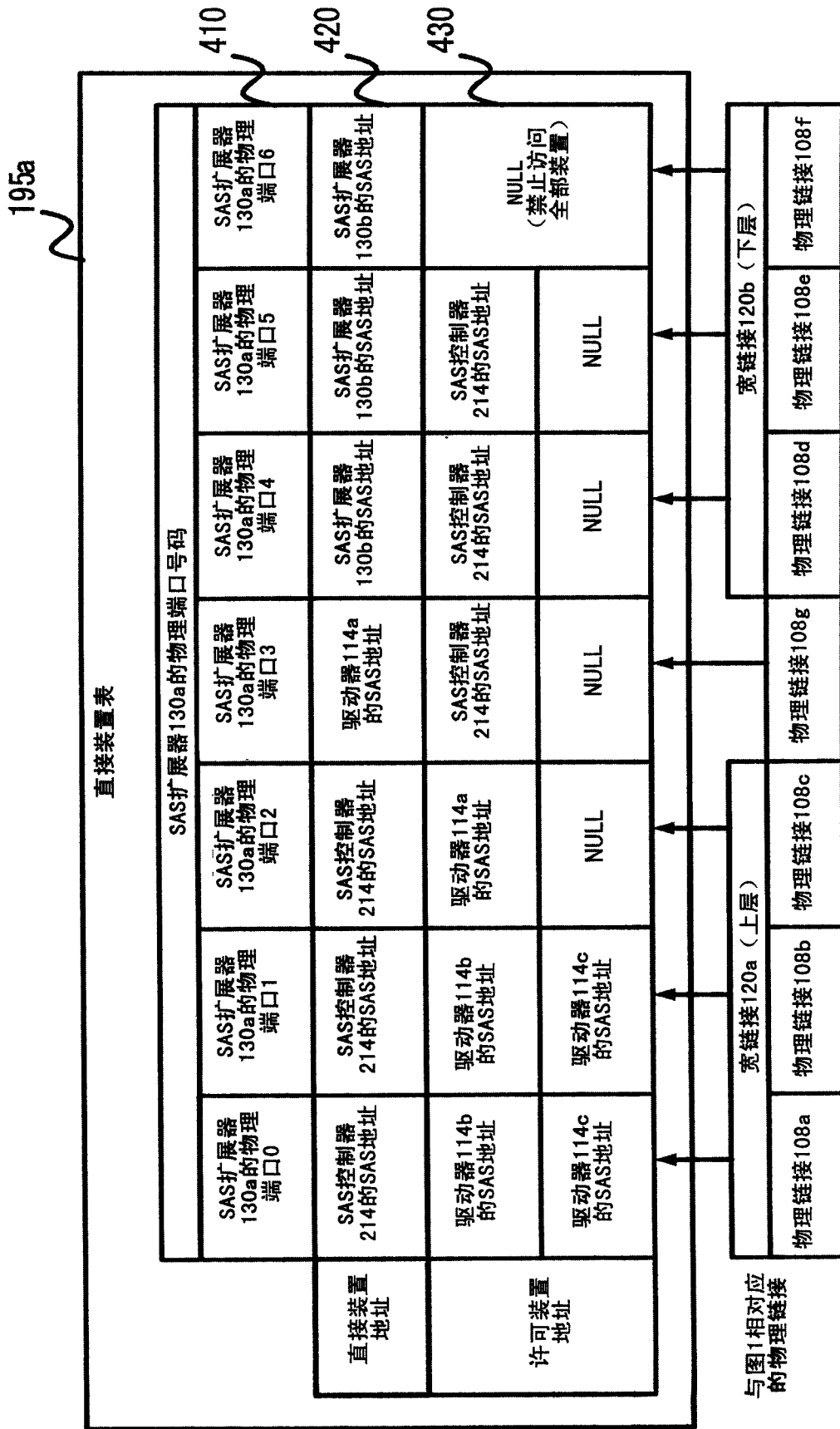


图 4



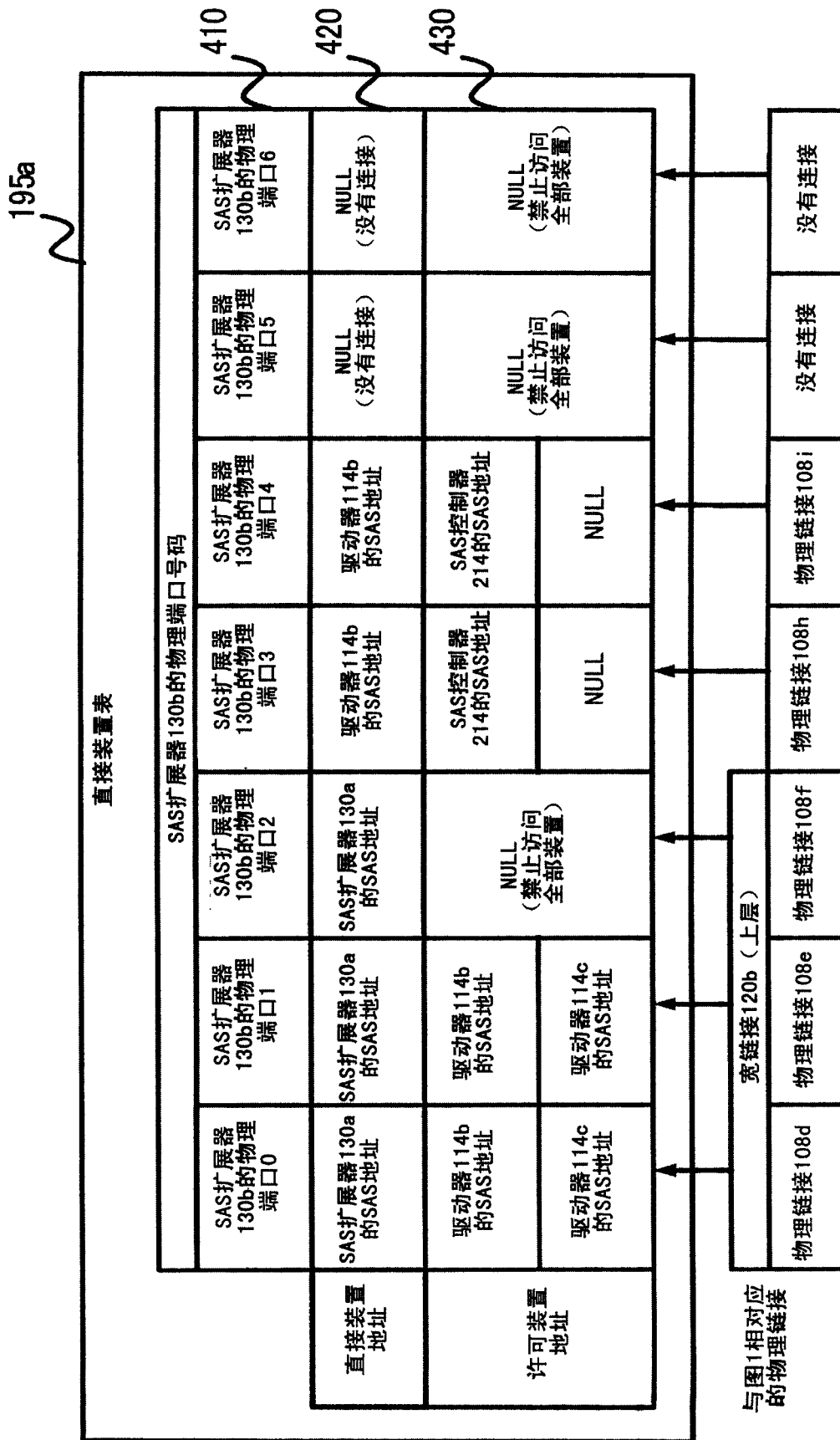


图 5

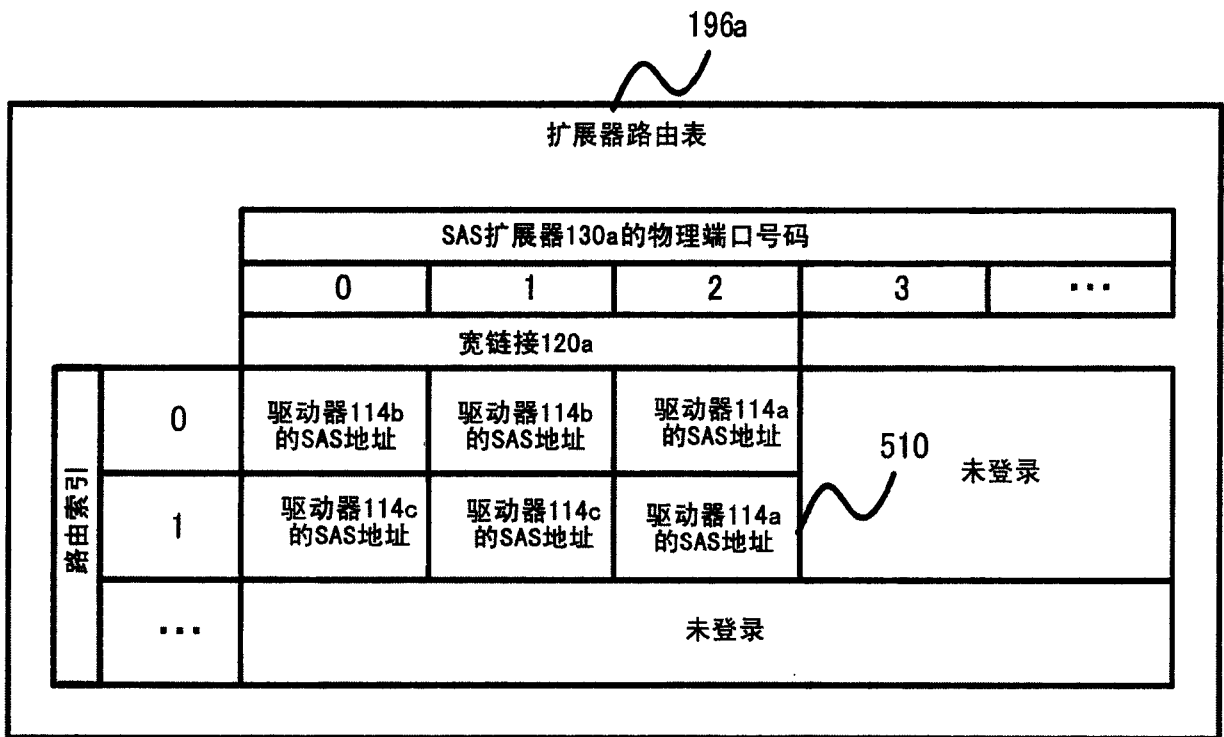


图 6

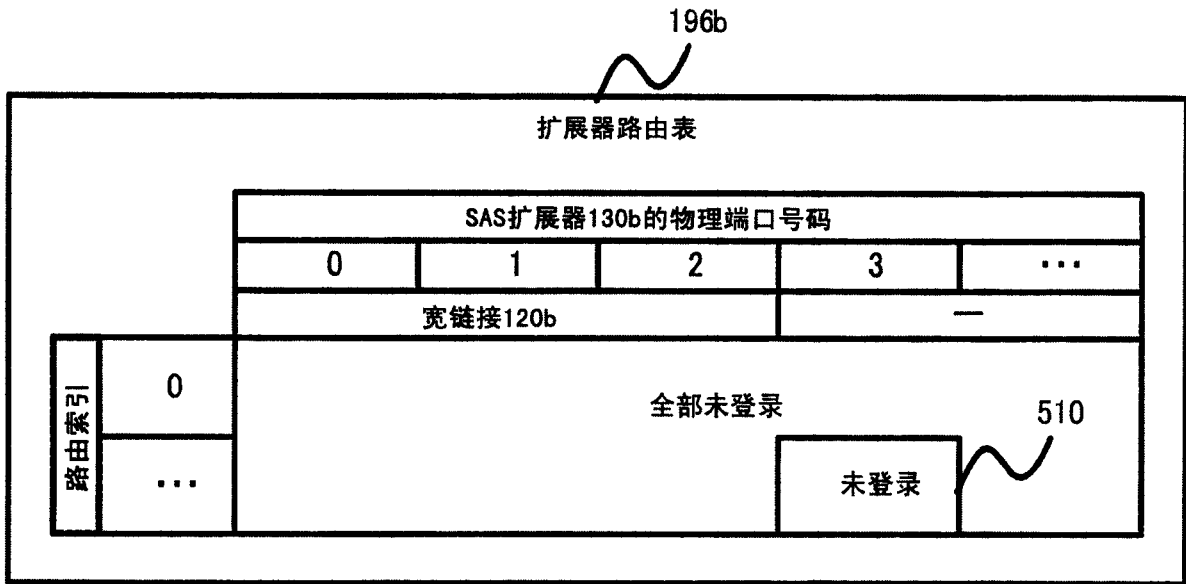


图 7

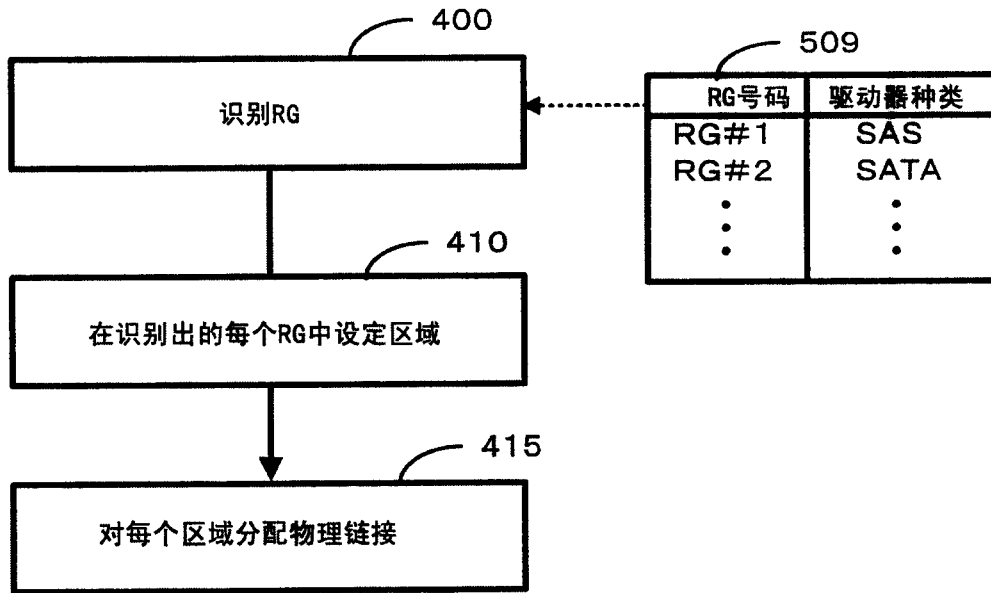


图 8

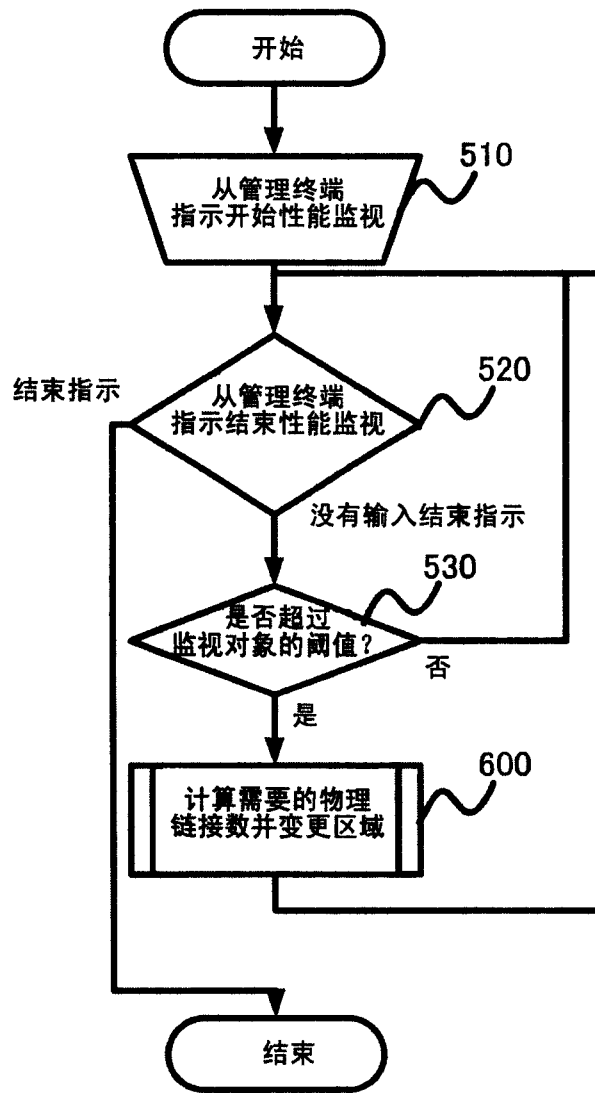


图 9

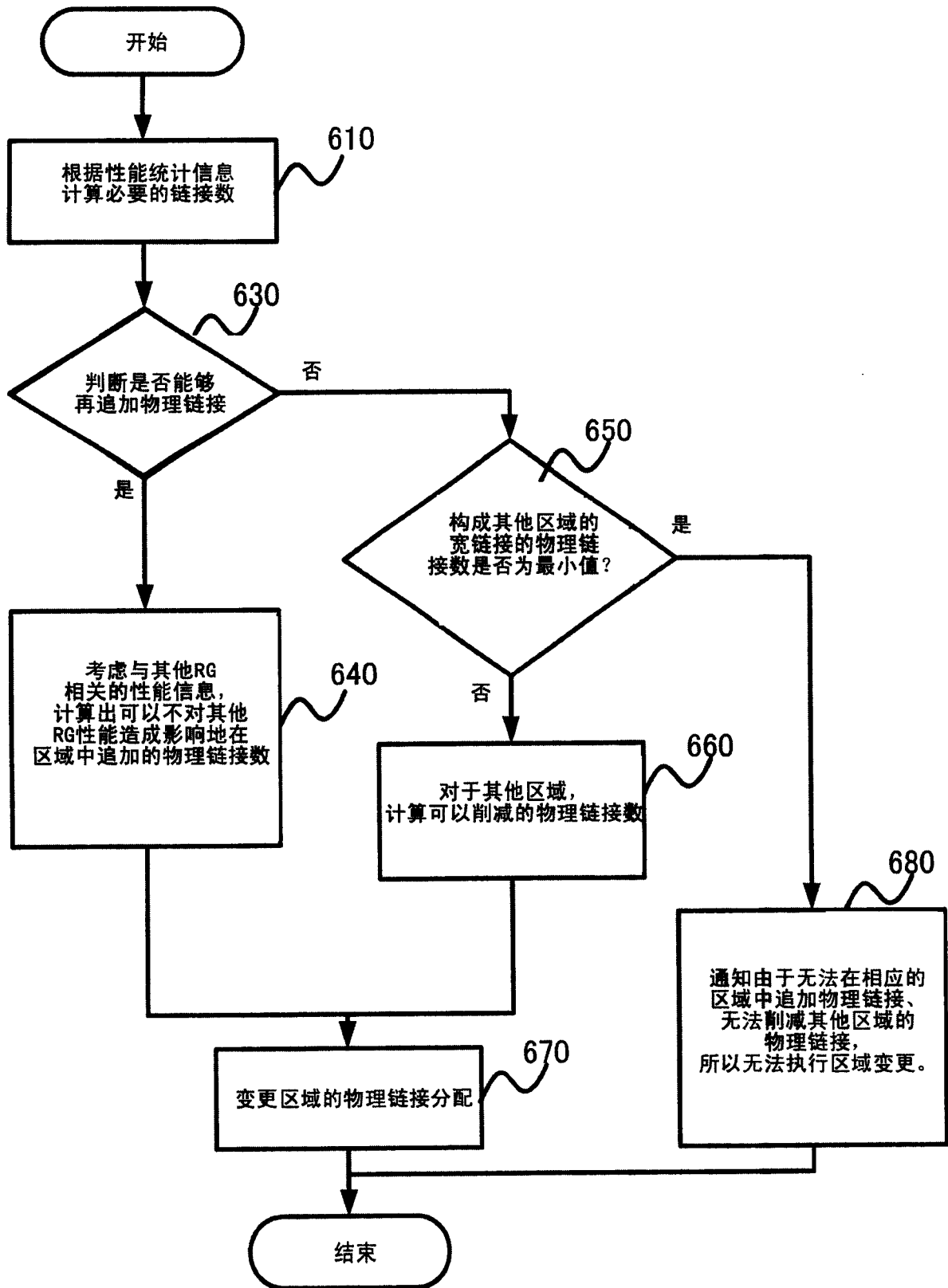


图 10

710	720	730	740	750	760	770	780	790
RG号码	驱动器种类	LU号码	区域号码	性能的统计信息	性能的阈值	构成WL的物理端口号码	宽链接发生局部故障	重叠的区域
RG#1	高性能SAS	LU#1 LU#3	区域#1	Time10: 700MB/s Time11: 700MB/s	超过 600MB/s 阈值	扩展器#1のphy#3 扩展器#1のphy#4	未发生	没有
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
RG#10	大容量、 低性能SATA	LU#10 LU#15 LU#20	区域#10	Time10: 70MB/s Time11: 70MB/s	150MB/s	扩展器#1のphy#1	未发生	区域#20
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
RG#20	大容量SAS	LU#40 LU#50	区域#10	Time10: 320MB/s Time11: 320MB/s	600MB/s	扩展器#1のphy#1 扩展器#1のphy#2	未发生	区域#10

图 11

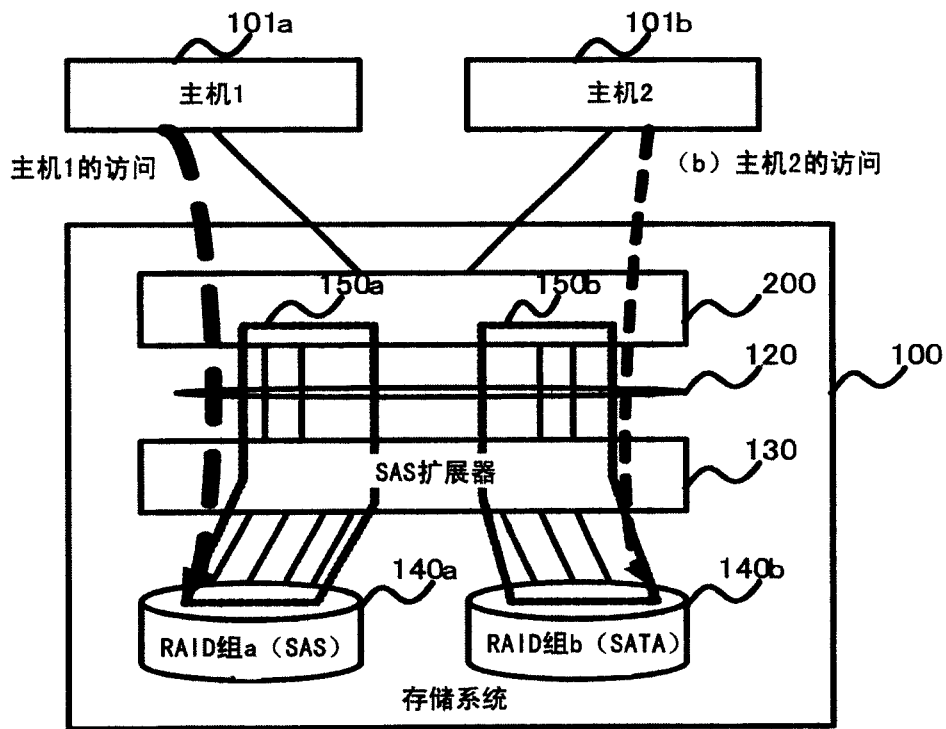


图 12A

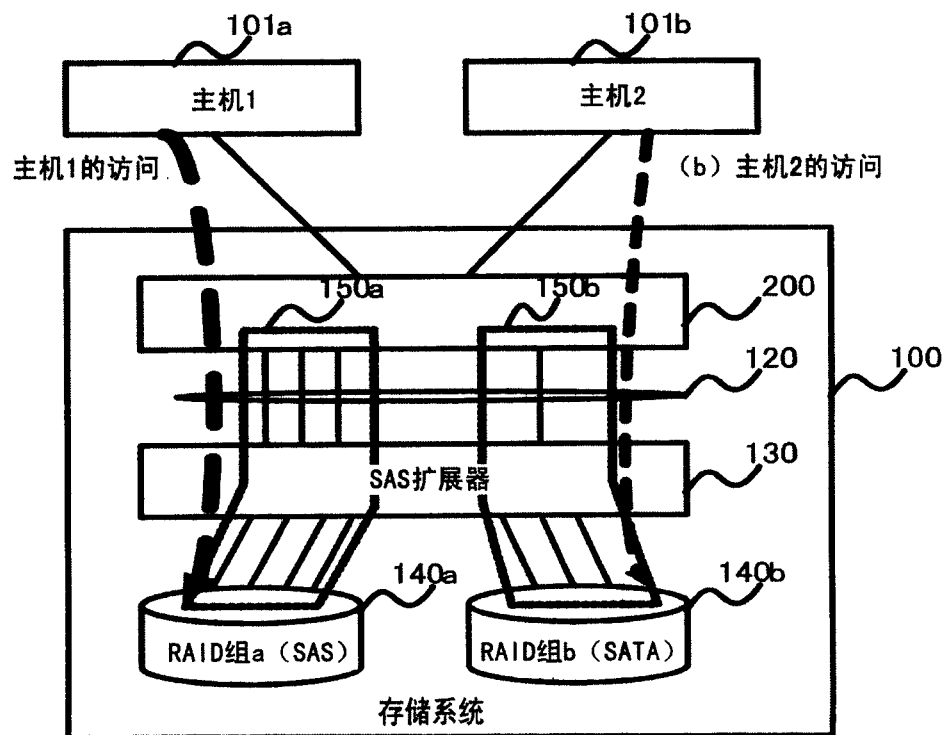


图 12B





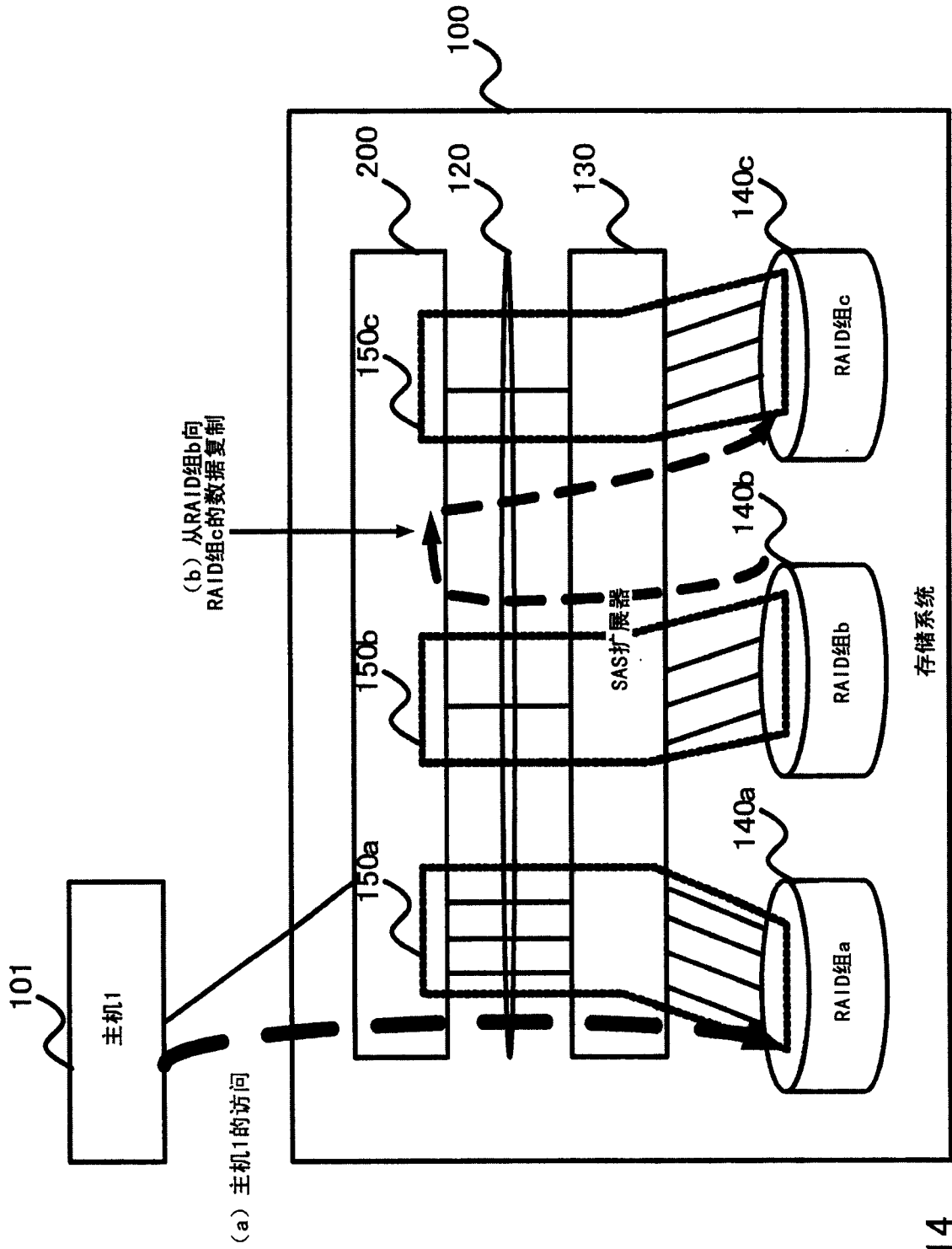


图 14

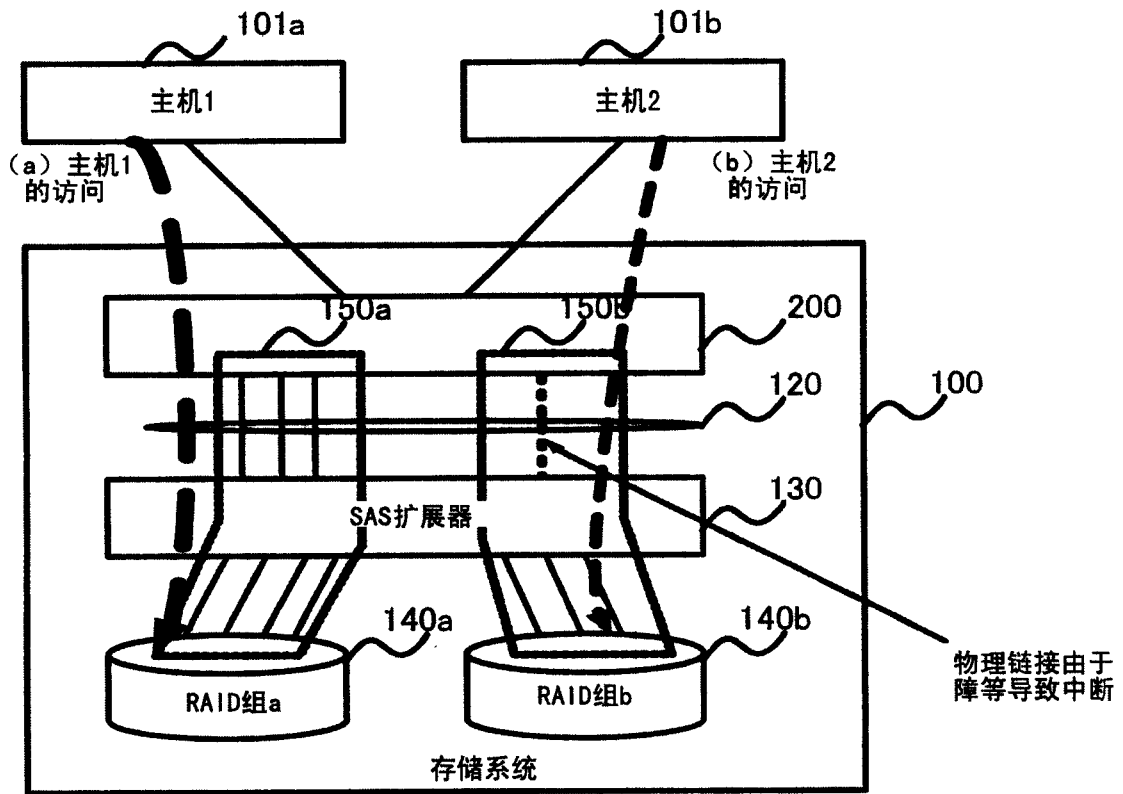


图 15A

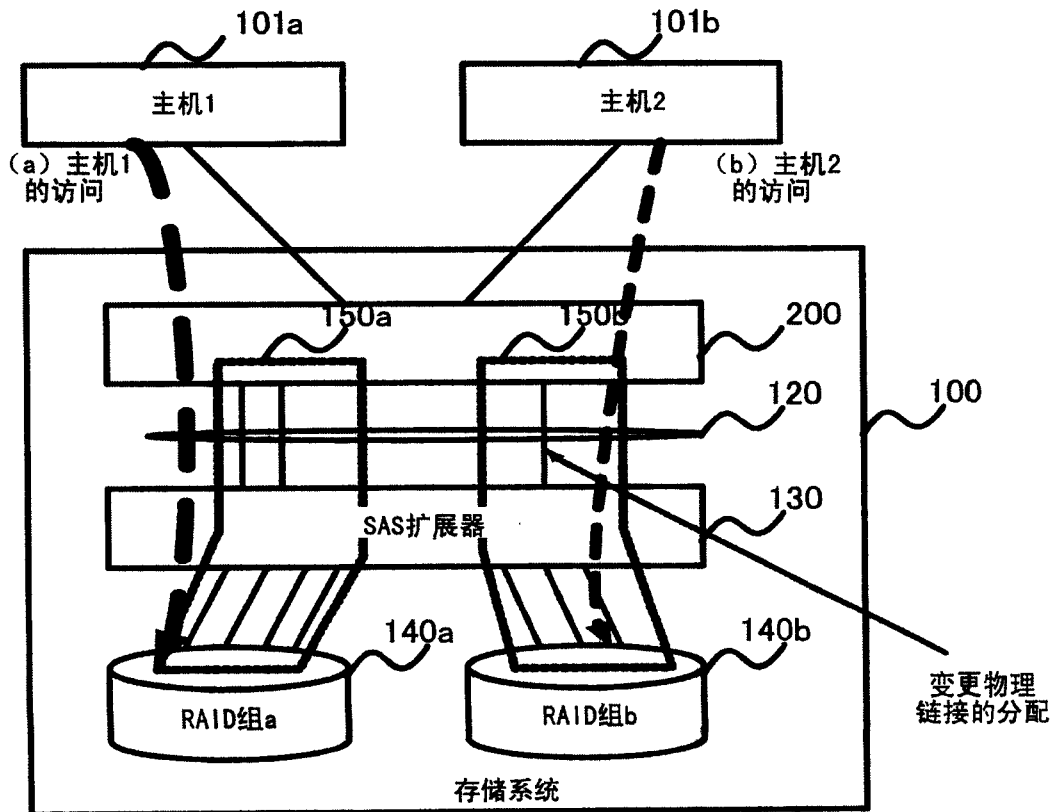


图 15B

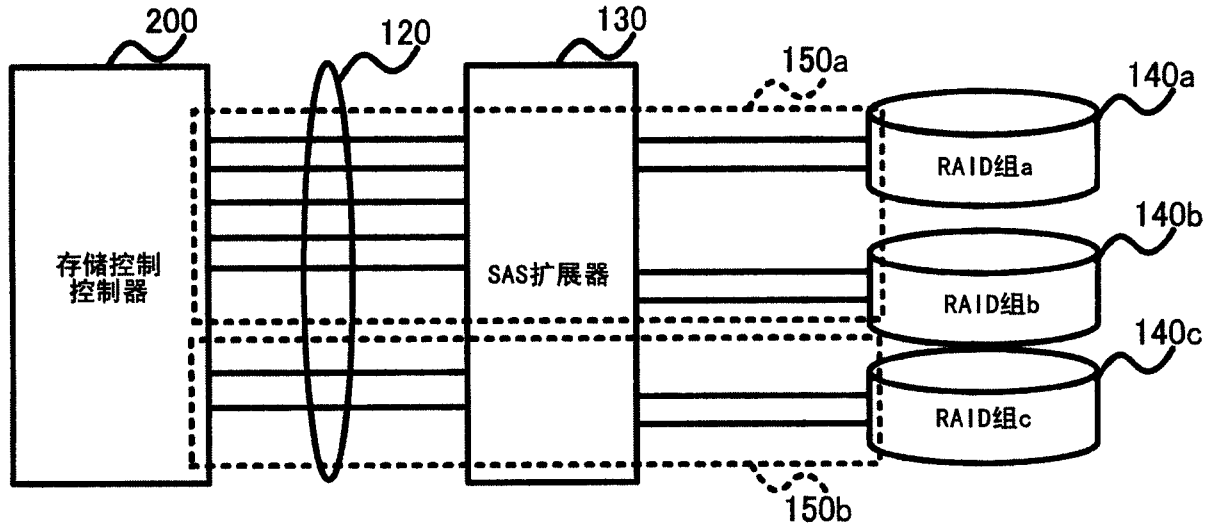


图 16

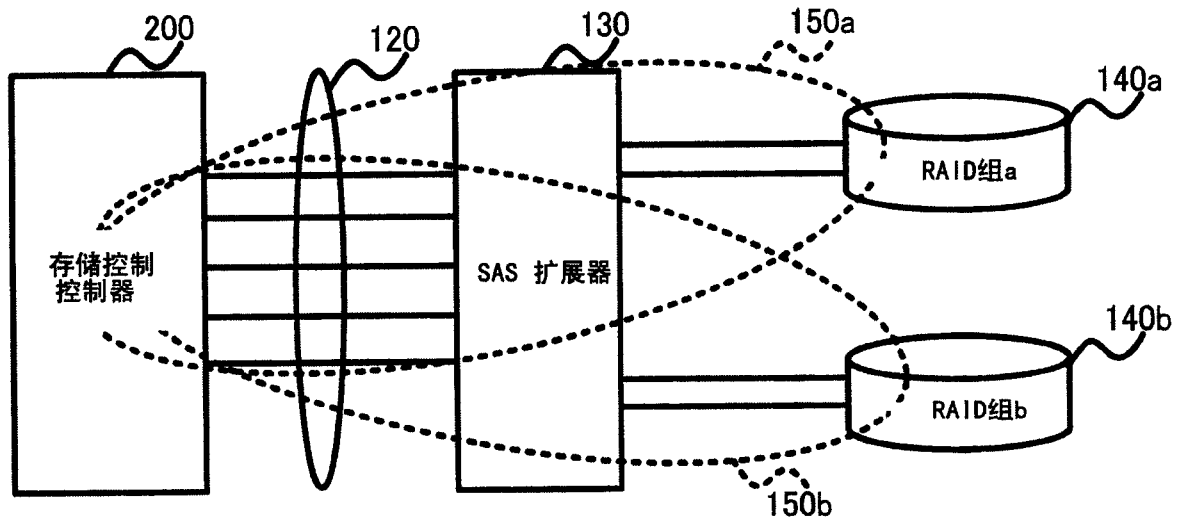


图 17

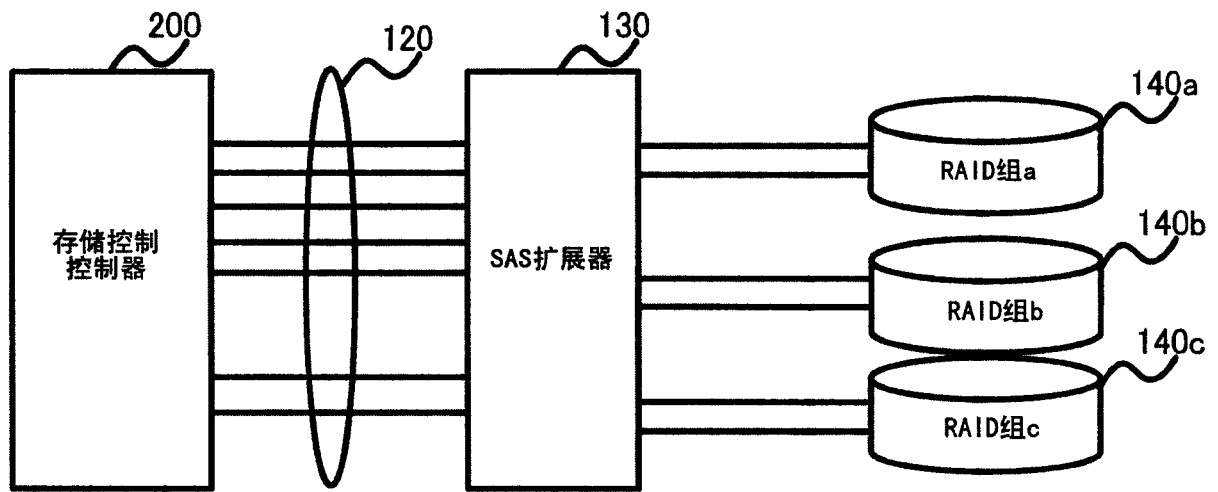


图 18

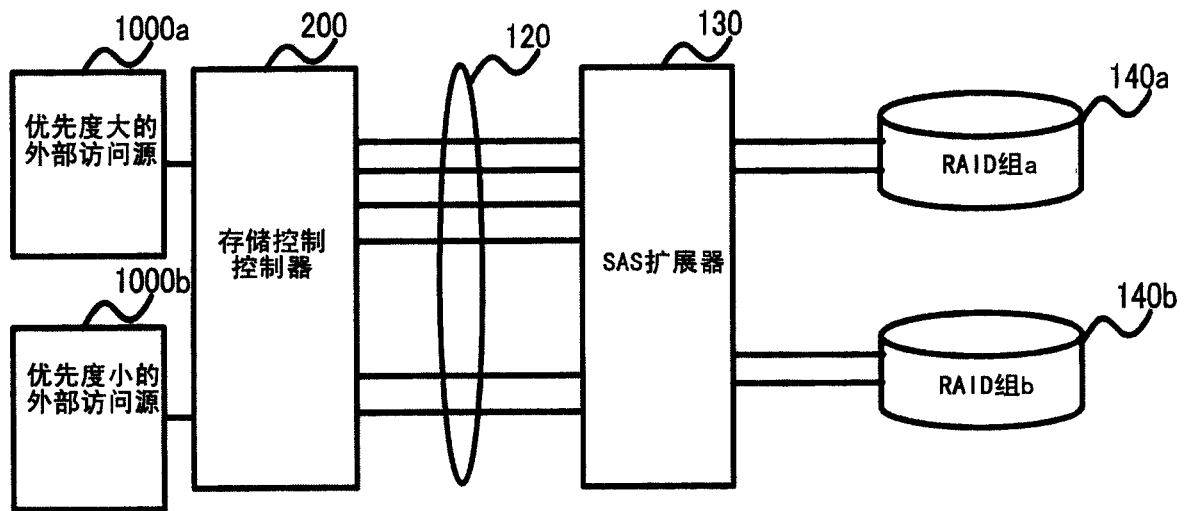


图 19

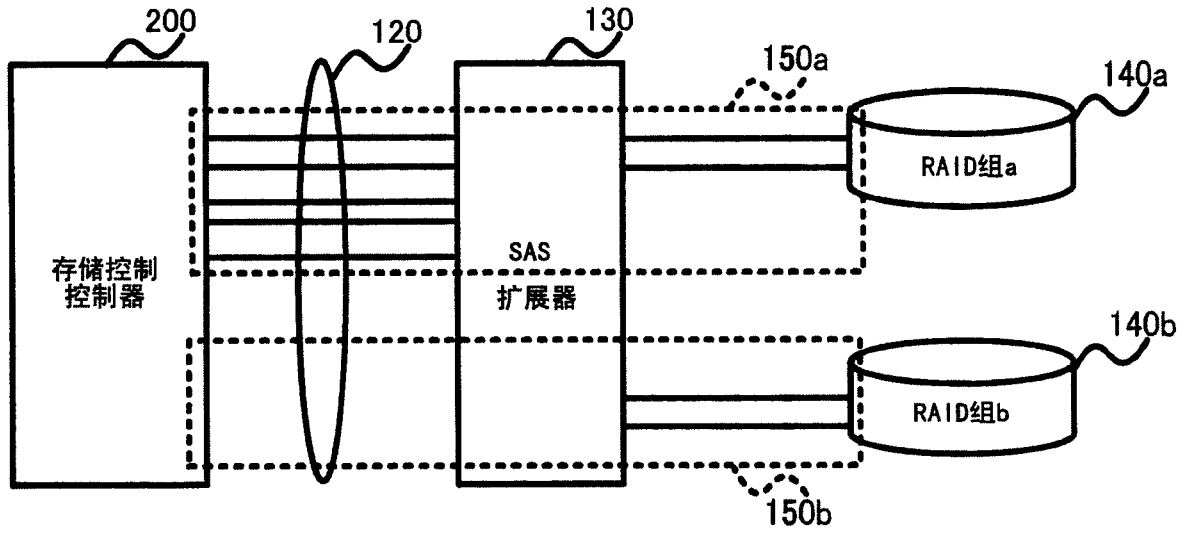


图 20



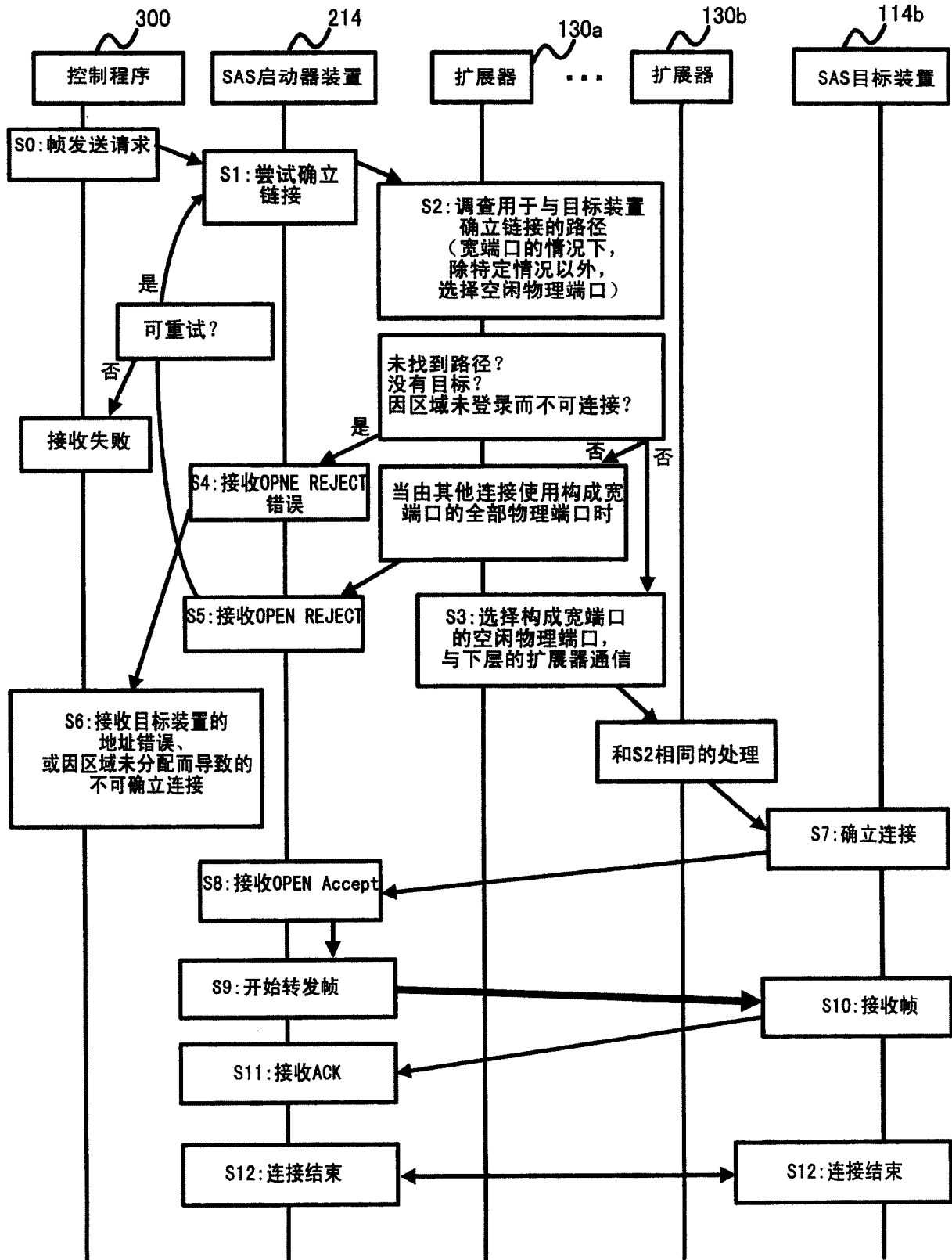


图 21

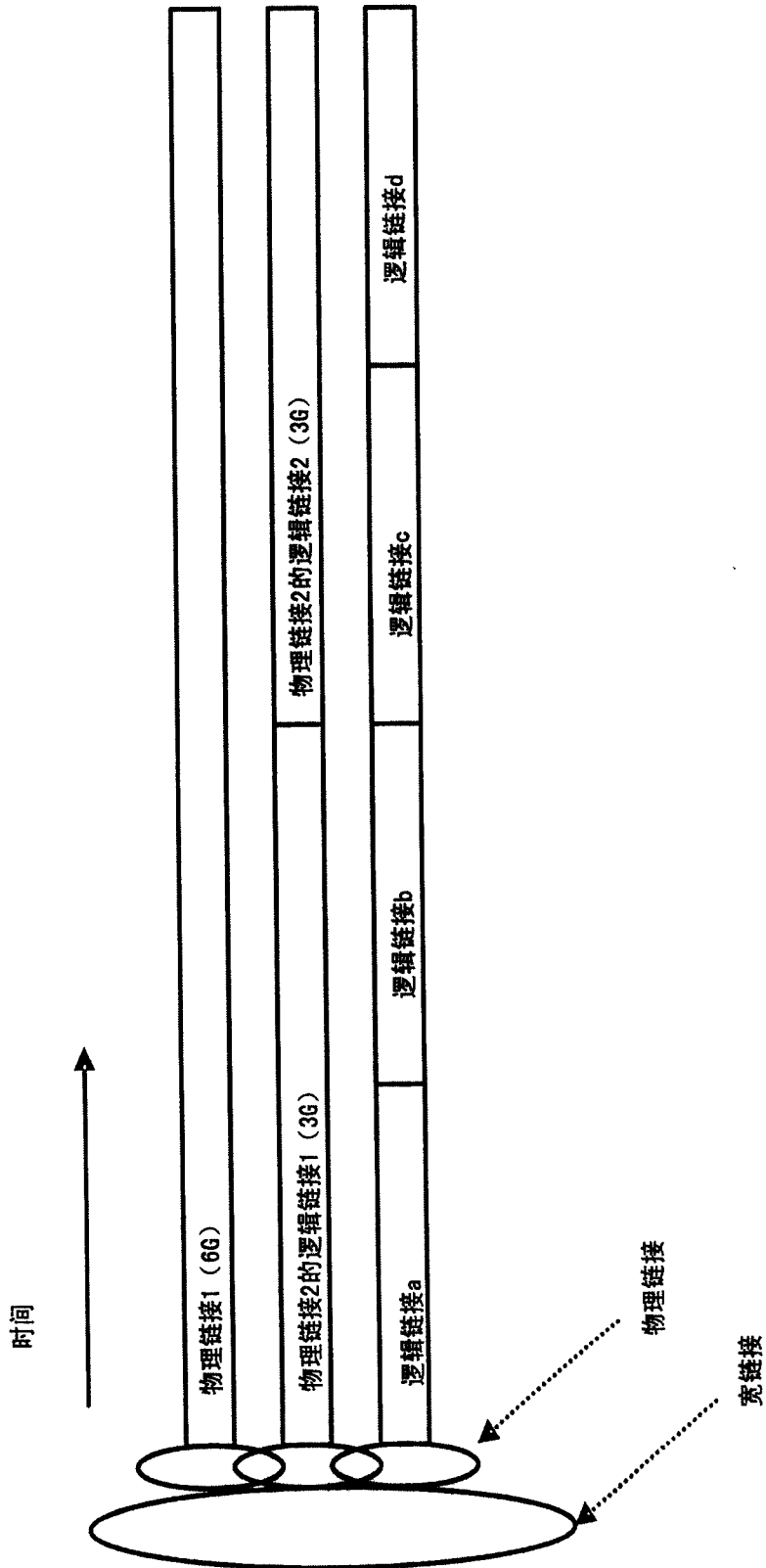


图 22