



## [12] 发明专利说明书

专利号 ZL 200610058636.7

[45] 授权公告日 2008 年 12 月 24 日

[11] 授权公告号 CN 100445966C

[22] 申请日 2006.3.2

[21] 申请号 200610058636.7

[30] 优先权

[32] 2005.9.5 [33] JP [31] 2005-255826

[73] 专利权人 株式会社日立制作所

地址 日本东京都

[72] 发明人 山本康友 本间久雄 里山爱

[56] 参考文献

US2003/0033494A1 2003.2.13

US2002/0184463A1 2002.12.5

US6240040B1 2001.5.29

CN1453715A 2003.11.5

US2004/0257857A1 2004.12.23

US6529976B1 2003.3.4

审查员 林萍娟

[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

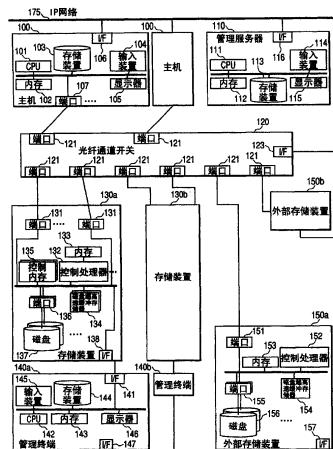
权利要求书 12 页 说明书 28 页 附图 14 页

[54] 发明名称

存储装置虚拟化装置的设备控制交接方法

[57] 摘要

在虚拟存储系统中，解决由于构成变更/故障发生/负载变动等引起的、提供虚拟设备的多个控制装置之间的负载的不均衡。在第一和第二存储装置连接第三存储装置的系统中，第一存储装置将第三存储装置的设备作为自己存储装置的设备虚拟化后来提供的过程中，将第一存储装置的超高速缓冲存储器保存的更新数据全部写入到第三存储装置的设备来反映，将该设备的属性交给第二存储装置，第二存储装置将该第三存储装置的设备作为自己存储装置的设备来进行虚拟化。



1. 一种计算机系统，其特征在于，

其具有：

第一计算机，

连接到所述第一计算机的第二计算机，

连接到所述第一以及所述第二计算机的第一存储装置系统，

连接到所述第一以及所述第二计算机的第二存储装置系统，

连接到所述第一以及所述第二存储装置系统的第三存储装置系统；

所述第一、第二以及第三存储装置系统具有存储所述第二计算机使用的数据的存储区域；

所述第一以及第二存储装置系统分别具有将连接到所述第一以及第二存储装置系统的第三存储装置系统具有的所述存储区域，作为所述第一以及第二存储装置系统具有的所述存储区域提供给所述第二计算机的虚拟化单元；

在所述第一存储装置系统的所述虚拟化单元，将所述第三存储装置系统内的第三存储区域作为所述第一存储装置系统内的第一存储区域提供给所述第二计算机的情况下，

所述第一计算机指示所述第二存储装置系统进行所述第二存储装置系统的访问权的变更，使其能够访问所述第三存储区域，

所述第一计算机指示所述第二存储装置系统，设定与所述第三存储区域对应的第二存储区域，

所述第二存储装置系统，根据所述设定指示设定所述第二存储区域，

所述第一计算机指示所述第二计算机识别所述第二存储区域，

所述第一计算机指示所述第二计算机，作为从所述第二计算机通过所述第一存储区域向所述第三存储区域进行输入输出请求的第一路径的交替，设定从所述第二计算机通过所述第二存储区域向所述第三存储区域进行输入输出请求的第二路径，

所述第一计算机指示所述第二计算机，禁止通过所述第一路径的对所述第三存储区域的输入输出请求，

---

所述第二存储装置系统的所述虚拟化单元，将所述第三存储区域作为所述第二存储区域提供给所述第二计算机。

2. 一种计算机系统，其特征在于，

其具有：

第一计算机，

连接到所述第一计算机的第二计算机，

连接到所述第一以及所述第二计算机的第一存储装置系统，

连接到所述第一以及所述第二计算机的第二存储装置系统，

连接到所述第一以及所述第二存储装置系统的第三存储装置系统；

所述第一、第二以及第三存储装置系统具有存储所述第二计算机使用的数据的存储区域；

所述第一以及第二存储装置系统分别具有将连接到所述第一以及第二存储装置系统的第三存储装置系统具有的所述存储区域，作为所述第一以及第二存储装置系统具有的所述存储区域提供给所述第二计算机的虚拟化单元；

所述第一存储装置系统的所述虚拟化单元，将所述第三存储装置系统内的第三存储区域作为所述第一存储装置系统内的第一存储区域提供给所述第二计算机的情况下，

所述第一计算机指示所述第二存储装置系统，设定与所述第一存储区域对应的第二存储区域；

所述第二存储装置系统根据所述设定指示设定所述第二存储区域；

所述第一计算机指示所述第二计算机识别所述第二存储区域；

所述第一计算机指示所述第二计算机，作为从所述第二计算机通过所述第一存储区域向所述第三存储区域进行输入输出请求的第一路径的交替，设定从所述第二计算机通过所述第二存储区域以及所述第一存储区域向所述第三存储区域进行输入输出请求的第三路径；

所述第一计算机指示所述第二计算机，禁止通过所述第一路径的向所述第三存储区域的输入输出请求；

所述第一计算机指示所述第二存储装置系统，进行访问权的变更，以使能够访问所述第三存储区域；

所述第一计算机指示所述第二存储装置系统变更对应关系，以使所述第三存储区域和所述第二存储区域对应起来；

所述第二存储装置系统的所述虚拟化单元，将所述第三存储区域作为所述第二存储区域提供给所述第二计算机。

3. 根据权利要求 1 所述的计算机系统，其特征在于，

所述第二计算机在从所述第一计算机接收到所述第二存储区域的识别指示的情况下，对于与分配给所述第二存储区域的第二逻辑设备对应的逻辑单元制作设备文件；

所述第二计算机，在从所述第一计算机作为所述第一路径的交替接收到所述第二路径的设定指示的情况下，分别与所述第一以及第二存储区域对应的所述设备文件作为与同一个存储区域对应的所述设备文件来识别。

4. 根据权利要求 1 所述的计算机系统，其特征在于，

所述第二计算机，在从所述第一计算机作为所述第一路径的交替接收所述第二路径的设定指示、接收到通过所述第一路径的对所述第三存储区域的输入输出请求的禁止指示的情况下，通过所述第二路径进行对所述第三存储区域的输入输出请求。

5. 根据权利要求 1 或 2 所述的计算机系统，其特征在于，

所述第一计算机取得所述第一存储装置系统的输入输出处理量、以及对所述第一存储装置系统的来自所述第二计算机的输入输出请求频度；

在所述输入输出处理量未达到用于判断存储器之间的 I/O 负载的差是否为规定水平以上的阈值的情况下、或所述输入输出请求频度为所述阈值及其以上的情况下，将所述第一存储装置系统具有的存储区域移动至所述第二存储装置系统。

6. 根据权利要求 5 所述的计算机系统，其特征在于，

在将所述第一存储装置系统具有的存储区域移动至所述第二存储装置系统的情况下，所述第二存储装置系统将提供了所述第三存储区域的所述第一存储区域作为所述第二存储区域来提供，由此将所述第三存储区域作为所述第二存储区域提供给所述第二计算机。

7. 根据权利要求 1 所述的计算机系统，其特征在于，

所述第一以及第二存储装置系统分别具有存储由所述第二计算机发送的数据的第一以及第二超高速缓冲存储器；

所述第一计算机指示所述第二存储装置系统设定第一模式，该第一模式为：在设定所述第二存储区域时，将存储在所述第二超高速缓冲存储器的、向所述第三存储区域发送的数据存储在所述第三存储区域后，对所述第二计算机发送存储处理结束信息；

所述第一计算机，在对所述第二计算机进行所述第二存储区域的识别指示后，指示所述第一存储装置系统设定第二模式，该第二模式为：将存储在所述第一超高速缓冲存储器的、向所述第三存储区域发送的数据存储在所述第三存储区域后，对所述第二计算机发送存储处理结束信息；

所述第一计算机对所述第二计算机进行所述第二存储区域的识别指示后，指示所述第一存储装置系统设定所述第二模式；

所述第一计算机，对所述第二计算机进行通过所述第一路径的对所述第三存储区域的输入输出请求的禁止指示后，指示解除所述第一模式设定；

所述第二存储装置系统，将从所述第二计算机发送到所述第三存储区域的数据存储到所述第二超高速缓冲存储器后，对所述第二计算机进行发送存储处理结束信息的设定。

#### 8. 根据权利要求 2 所述的计算机系统，其特征在于，

所述第一以及第二存储装置系统分别具有存储从所述第二计算机发送的数据的第一以及第二超高速缓冲存储器；

所述第一计算机对所述第二计算机进行所述第二存储区域的识别指示后，指示所述第二存储装置系统设定第一模式，该第一模式为：将存储在所述第二超高速缓冲存储器的、向所述第三存储区域发送的数据存储到所述第三存储区域后，对所述第二计算机发送存储处理结束信息；

所述第一计算机对所述第二计算机进行通过所述第一路径的、对所述第三存储区域的输入输出请求的禁止指示后，指示对所述第二存储装置系统的所述第一模式设定的解除；

所述第一计算机指示所述第一存储装置系统设定第二模式，该第二模式为：将存储在所述第一超高速缓冲存储器的、向所述第三存储区域发送的数

据存储到所述第三存储区域后，对所述第二计算机发送存储处理结束信息；

所述第一计算机对所述第二存储装置系统变更访问权，以使能够访问所述第三存储区域。

9. 根据权利要求 7 所述的计算机系统，其特征在于，

所述第一存储装置系统，在从所述第二计算机对所述第三存储区域有输入输出请求的情况下，对所述第三存储区域判断是否允许了来自所述第二计算机的访问的访问权；

在没有允许所述访问的情况下，拒绝来自所述第二计算机的输入输出请求；

在允许了所述访问的情况下，判断是否设定了所述第二模式；

在没有设定所述第二模式的情况下，对来自所述第二计算机的输入输出请求执行输入输出处理；

在设定了所述第二模式的情况下，对来自所述第二计算机的输出请求执行输出处理；

对来自所述第二计算机的输入请求，将由所述第二计算机发送的数据存储到所述第一超高速缓冲存储器，在存储到所述第三存储区域后，对所述第二计算机发送所述输入处理的结束信息。

10. 一种存储装置系统，其特征在于，

具有控制部、存储装置、超高速缓冲存储器以及通信部；

所述控制部、所述存储装置、所述超高速缓冲存储器和所述通信部，通过相互连接网连接；

所述通信部被连接到第一计算机、第二计算机、连接到所述第一计算机的第一存储装置系统、连接到所述第一存储装置系统的第二存储装置系统；

在所述第一存储装置系统将所述第二存储装置系统具有的第二存储区域作为所述第一存储装置系统内的第一存储区域提供给所述第一计算机的情况下，

所述控制部，在所述存储装置设定与所述第二存储区域对应的第三存储区域，

所述控制部从所述第一存储装置系统接收所述第一存储区域的识别信息

后，在所述第三存储区域设定；

在所述第二计算机，指示所述第一计算机识别所述第三存储区域，作为通过所述第一存储区域进行对所述第二存储区域的输入输出请求的路径的交替，设定通过所述第三存储区域进行对所述第二存储区域的输入输出请求的路径的情况下，

所述控制部将所述第二存储区域作为所述第三存储区域提供给所述第一计算机。

11. 一种存储装置系统，其特征在于，

具有控制部、存储装置、超高速缓冲存储器以及通信部；

所述控制部、所述存储装置、所述超高速缓冲存储器和所述通信部，通过相互连接网连接；

所述通信部被连接到第一计算机、第二计算机、连接到所述第一计算机的第一存储装置系统、连接到所述第一存储装置系统的第二存储装置系统；

所述第一存储装置系统，将所述第二存储装置系统具有的第二存储区域作为所述第一存储装置系统内的第一存储区域提供给所述第一计算机的情况下，

所述控制部在所述存储装置设定与所述第一存储区域对应的第三存储区域，

所述控制部从所述第一存储装置接收所述第一存储区域的识别信息后，在所述第三存储区域设定；

所述第二计算机，指示所述第一计算机识别所述第三存储区域，作为通过所述第一存储区域进行对所述第二存储区域的输入输出请求的路径的交替，设定通过所述第三存储区域和所述第一存储区域进行对所述第二存储区域的输入输出请求的路径的情况下，

所述控制部变更对应关系，以使所述第二存储区域和所述第三存储区域对应起来，

所述控制部将所述第二存储区域作为所述第三存储区域提供给所述第一计算机。

12. 根据权利要求 10 所述的存储装置系统，其特征在于，

所述控制部在所述存储装置设定与所述第一存储区域对应的第三存储区域时，设定一种处理模式，该处理模式为：将存储在所述超高速缓冲存储器的、向所述第二存储区域发送的数据存储到所述第二存储区域后，对所述第一计算机发送存储处理结束信息；

在所述第二计算机，指示所述第一计算机识别所述第三存储区域，作为通过所述第一存储区域进行对所述第二存储区域的输入输出请求的路径的交替，设定通过所述第三存储区域进行对所述第二存储区域的输入输出请求的路径的情况下，

所述控制部，解除所述处理模式，将从所述第一计算机发送到所述第二存储区域的数据存储到所述超高速缓冲存储器后，对所述第一计算机发送存储处理结束信息。

13. 根据权利要求 10 或 11 所述的存储装置系统，其特征在于，

在从所述第一计算机对所述第二存储区域有输入输出请求的情况下，

所述控制部对所述第二存储区域判断是否允许了来自所述第一计算机的访问；

在没有允许所述访问的情况下，拒绝来自所述第一计算机的输入输出请求；

在允许了所述访问的情况下，判断是否设定了在所述第二存储区域中进行了存储之后，对所述第一计算机发送存储处理结束信息的处理模式；

在没有设定所述处理模式的情况下，对来自所述第一计算机的输入输出请求执行输入输出处理；

在设定了所述处理模式的情况下，对来自所述第一计算机的输出请求执行输出处理；

对来自所述第一计算机的输入请求，将从所述第一计算机对所述第二存储区域发送的数据存储到所述超高速缓冲存储器，将存储在超高速缓冲存储器的数据存储到所述第二存储区域后，向所述第一计算机发送所述输入处理已结束的信息。

14. 一种计算机系统中的访问控制方法，其特征在于，

该计算机系统具有：

第一计算机，

连接到所述第一计算机的第二计算机，

连接到所述第一以及所述第二计算机的第一存储装置系统，

连接到所述第一以及所述第二计算机的第二存储装置系统，

连接到所述第一以及所述第二存储装置系统的第三存储装置系统；

所述第一、第二以及第三存储装置系统分别具有存储所述第二计算机使用的数据的存储区域；

在所述第一存储装置系统将所述第三存储装置系统内的第三存储区域作为所述第一存储装置系统内的第一存储区域提供给所述第二计算机的情况下，

变更访问权，以使从所述第二存储装置能够访问所述第三存储区域，

在所述第二存储装置系统设定与所述第三存储区域对应的第二存储区域，

对所述第二存储区域设定所述第一存储区域的识别信息，

所述第二计算机识别所述第二存储区域，

作为从所述第二计算机通过所述第一存储区域对所述第三存储区域进行输入输出请求的第一路径的交替，设定从所述第二计算机通过所述第二存储区域对所述第三存储区域进行输入输出请求的第二路径，

禁止从所述第二计算机通过所述第一路径的输入输出请求，

所述第二存储装置系统将所述第三存储区域作为所述第二存储区域提供给所述第二计算机。

15. 一种计算机系统中的访问控制方法，其特征在于，

该计算机系统具有：

第一计算机，

连接到所述第一计算机的第二计算机，

连接到所述第一以及所述第二计算机的第一存储装置系统，

连接到所述第一以及所述第二计算机的第二存储装置系统，

连接到所述第一以及所述第二存储装置系统的第三存储装置系统；

所述第一、第二以及第三存储装置系统分别具有存储所述第二计算机使

用的数据的存储区域；

在所述第一存储装置系统将所述第三存储装置系统内的第三存储区域作为所述第一存储装置系统内的第一存储区域提供给所述第二计算机的情况下，

在所述第二存储装置系统设定与所述第一存储区域对应的第二存储区域；

对所述第二存储区域设定所述第一存储区域的识别信息；

所述第二计算机识别所述第二存储区域；

作为从所述第二计算机通过所述第一存储区域对所述第三存储区域进行输入输出请求的第一路径的交替，设定从所述第二计算机通过所述第二存储区域以及所述第一存储区域对所述第三存储区域进行输入输出请求的第三路径；

禁止从所述第二计算机通过所述第一路径的输入输出请求；

变更访问权，以使从所述第二存储装置系统能够访问所述第三存储区域；

变更对应关系，以使所述第三存储区域和所述第二存储区域对应起来；

所述第二存储装置系统将所述第三存储区域作为所述第二存储区域提供给所述第二计算机。

16. 根据权利要求 14 所述的访问控制方法，其特征在于，

所述第二计算机识别所述第二存储区域时，对于与分配给所述第二存储区域的第二逻辑设备对应的逻辑单元，制作设备文件；

作为从所述第二计算机通过所述第一存储区域对所述第三存储区域进行输入输出请求的第一路径的交替，设定从所述第二计算机通过所述第二存储区域对所述第三存储区域进行输入输出请求的第二路径时，分别与所述第一以及第二存储区域对应的所述设备文件作为与同一个存储区域对应的所述设备文件来识别。

17. 根据权利要求 14 所述的访问控制方法，其特征在于，

在禁止从所述第二计算机通过所述第一路径的输入输出请求的情况下，通过所述第二路径进行从所述第二计算机对所述第三存储区域的输入输出请求。

18. 根据权利要求 14 或 15 所述的访问控制方法，其特征在于，  
检测所述第一存储装置系统的输入输出处理量，以及对所述第一存储装  
置系统的、来自所述第二计算机的输入输出请求频度；

在所述输入输出处理量不超过用于判断存储器之间的 I/O 负载的差是否  
为规定水平以上的阈值的情况下、或所述输入输出请求频度为所述阈值或其  
以上的情况下，进行交替所述第一存储区域，将所述第三存储区域作为所述  
第二存储区域提供给所述第二计算机的处理。

19. 根据权利要求 14 所述的访问控制方法，其特征在于，

在该计算机系统中，所述第一以及第二存储装置系统分别具有存储由所  
述第二计算机发送的数据的第一以及第二超高速缓冲存储器；

在所述第二存储装置系统设定所述第二存储区域时，设定第一模式，该  
第一模式为：将存储在所述第二超高速缓冲存储器中的、向所述第三存储区  
域发送的数据存储到所述第三存储区域后，对所述第二计算机发送存储处理  
结束信息；

在所述第二计算机识别所述第二存储区域后，对所述第一存储装置系统  
设定第二模式，该第二模式为：将存储在所述第一超高速缓冲存储器中的、  
向所述第三存储区域发送的数据存储到所述第三存储区域后，对所述第二计  
算机发送存储处理结束信息；

禁止从所述第二计算机通过所述第一路径的、对所述第三存储区域的输  
入输出请求后，解除所述第一模式。

20. 根据权利要求 15 所述的访问控制方法，其特征在于，

在该计算机系统中，所述第一以及第二存储装置系统分别具有存储由所  
述第二计算机发送的数据的第一以及第二超高速缓冲存储器；

在所述第二计算机识别所述第二存储区域后，设定第一模式，该第一模  
式为：将存储在所述第二超高速缓冲存储器中的、向所述第三存储区域发  
送的数据存储到所述第三存储区域后，对所述第二计算机发送存储处理结束信  
息；

禁止从所述第二计算机通过所述第一路径的、对所述第三存储区域的输  
入输出请求后，对所述第二存储装置系统解除所述第一模式，对所述第一存

储装置系统，将存储在所述第一超高速缓冲存储器中的、向所述第三存储区域发送的数据存储到所述第三存储区域后，对所述第二计算机发送存储处理结束信息；

对所述第二存储装置系统变更访问权，以使能够访问所述第三存储区域。

21. 一种计算机，其特征在于，

具有控制部和通信部；

所述通信部连接到第一、第二以及第三存储装置系统；

所述第一存储装置系统将所述第三存储装置系统具有的第三存储区域作为所述第一存储装置系统内的第一存储区域提供给其他计算机的情况下，

所述控制部指示所述第二存储装置系统变更访问权，使其能够访问所述第三存储区域，

所述控制部指示所述第二存储装置系统设定与所述第三存储区域对应的第二存储区域，

所述控制部指示所述其他计算机识别所述第二存储区域，

所述控制部指示所述其他计算机，作为通过所述第一存储区域对所述第三存储区域进行输入输出请求的第一路径的交替，设定通过所述第二存储区域对所述第三存储区域进行输入输出请求的第二路径，

所述控制部指示所述其他计算机禁止通过所述第一路径的、对所述第三存储区域的输入输出请求。

22. 一种计算机，其特征在于，

具有控制部和通信部；

所述通信部连接到第一、第二以及第三存储装置系统；

在所述第一存储装置系统将所述第三存储装置系统内的第三存储区域作为所述第一存储装置系统内的第一存储区域提供给所述其他计算机的情况下，

所述控制部指示所述第二存储装置系统设定与所述第一存储区域对应的第二存储区域，

所述控制部指示所述其他计算机识别所述第二存储区域；

所述控制部指示所述其他计算机，作为通过所述第一存储区域对所述第三存储区域进行输入输出请求的第一路径的交替，设定通过所述第二存储区域和所述第一存储区域对所述第三存储区域进行输入输出请求的第三路径；

所述控制部指示所述其他计算机禁止通过所述第一路径的对所述第三存储区域的输入输出请求；

所述控制部指示所述第二存储装置系统变更访问权，以使能够访问所述第三存储区域；

所述控制部指示所述第二存储装置系统变更对应关系，以使所述第三存储区域和所述第二存储区域对应起来。

23. 根据权利要求 21 或 22 所述的计算机，其特征在于，

所述控制部，取得所述第一以及第二存储装置系统的构成信息以及动作信息，根据所取得的所述构成信息以及所述动作信息，发送将一方的所述存储装置系统具有的所述存储区域移动至他方的所述存储装置系统的移动指示。

24. 根据权利要求 23 所述的计算机，其特征在于，

所述控制部作为所述第一以及第二存储装置系统的所述构成信息取得所述各存储装置系统的输入输出处理量；

在所述各存储装置系统的一方的所述输入输出处理量不超过用于判断存储器之间的 I/O 负载的差是否为规定水平以上的阈值的情况下，根据所述移动指示，决定从该存储装置系统移动所述存储区域；

在所述各存储装置系统的所述输入输出处理量超过所述阈值的情况下，

所述控制部作为所述第一以及第二存储装置系统的所述动作信息取得对所述各存储装置系统的、来自所述其他计算机的输入输出请求频度；

在所述各存储装置系统的一方的所述输入输出请求频度为所述阈值和其以上的情况下，根据所述移动指示，决定从该存储装置系统移动所述存储区域。

25. 根据权利要求 23 所述的计算机，其特征在于，

在所述第一以及第二存储装置系统中一方的存储装置系统，将连接在该存储装置系统的所述第三存储系统的所述存储区域作为该存储装置系统内的第四存储区域提供给所述其他计算机的情况下，

所述控制部在发送所述移动指示时，优先选择所述第四存储区域作为移动对象。

## 存储装置虚拟化装置的设备控制交接方法

### 技术领域

本发明涉及在计算机系统中，将存储计算机使用的数据的存储装置系统（以下，称为存储装置）、特别是主机计算机（以下，称为主机）和一台或其以上的其他存储装置连接起来，将其他存储装置内的存储区域（以下，称为设备）作为自己的设备虚拟化的存储装置的控制方法。

### 背景技术

随着因特网商务的扩大、手续的电子化等信息系统的急速的发展，计算机收发的数据量飞跃地增加。除了这样的数据量的飞跃的增加，基于向磁盘装置的数据备份（Disk-to-Data Backup）或检查跟踪对应等的企业的业务活动记录（交易信息或邮件等）的长期保管等、存储在存储装置的数据量也在飞跃地增加。随之，在企业信息系统中，在谋求各部门/各系统的存储装置的增强的同时，还谋求复杂化的IT基础设施的管理的简化或高效率化。特别，对简化存储装置的管理、根据数据的价值活用最合适的存储装置来谋求总成本的最优化的技术的期待越来越高。

作为降低具有大规模存储装置的系统的管理成本的方法，有特开平2005-011277号公报中公开的存储装置虚拟化技术。在专利文献1中，公开了一种存储装置虚拟化技术（以下，称为外部存储装置连接技术）：即将第一存储装置连接到一台或其以上的第二存储装置，将第二存储装置向主机等上级装置提供的设备作为第一存储装置的逻辑设备（以下，称为逻辑设备），通过第一存储装置向主机提供。第一存储装置，在接收到来自主机的向逻辑设备的输入输出请求时，判断访问对象的设备是对应于第二存储装置的逻辑设备、或者是包含在第一存储装置的磁盘装置等物理设备（以下，称为物理设备）中的哪一个，根据判定结果对适当的访问目的地分配输入输出请求。

利用使用了特开平2005-011277号公报中所示的外部存储装置连接技术的存储装置，就能够构成合并了性能、可靠性、价格等属性不同的多个存

储装置的系统。例如，连接利用了外部存储装置连接技术的高成本、高功能、高可靠性的第一存储装置和多台低成本、低功能、低可靠性的第二存储装置，将存储区域分级，由此可以进行与数据的新鲜度或价值对应的最佳的数据配置。通过使用这样的具有存储阶层不同的存储装置的系统，可以以检查跟踪对应为目的，将每天的业务中发生的交易信息或邮件等大量的信息，根据各信息具有的价值用最佳的成本长期保存。

但是，在超过存储装置寿命的数十年程度的长期保存上述的大量的数据的情况下，在数据的保存期间内需要交换构成计算机系统的设备。特别是，在交换使用了外部存储装置连接技术的第一存储装置时，在专利文献1中，公开了将转移源的第一存储装置所管理的设备的数据复制后转移到成为新导入的数据的转移目的地的第一存储装置。通过该数据的复制处理，对原有的第一存储装置进行的设备的输入输出处理可能带来吞吐量低下等恶劣影响。再有，因为需要复制全部的成为转移对象的大量的设备的数据，因此第一存储装置之间的设备交接需要花很多的时间。

另外，在具有处理大规模数据中心等大量的数据的存储装置的系统中，根据确保性能等观点，还要预想在系统内连接多台第一存储装置来运用。在这样的系统的情况下，因为各个第一存储装置将第二存储装置所具有的大量的设备虚拟化，因此根据对各设备的输入输出负载（以下，将对设备的输入输出还称为I/O）的变动，第一存储装置之间的负载的不均匀就会扩大。

#### 发明内容

本发明的目的在于，在利用了存储装置虚拟化技术的计算机系统中，由于设备的追加或削减等系统结构的变动、装置故障引起的特定存储装置的输入输出处理能力下降、以及主机输入输出负载的变动等影响，重新调整存储装置处理能力和输入输出负载的平衡。

为了将存储装置之间的输入输出负载均匀化/最佳化，对于担当设备的虚拟化的存储装置（以下，称为虚拟化控制存储装置）之间的设备交接，使用以下的方法实现。

在由主机、管理服务器、使用了外部存储装置技术的第一和第二虚拟化控制存储装置（以下，还称为第一（第二）存储装置）、通过外部存储装置连

接来虚拟化的一台或其以上的第三存储装置构成的计算机系统中，第三存储装置内的第三设备作为第一存储装置的第一设备被虚拟化。如第三设备那样，以下，将从第一存储装置看是其他存储装置内的设备、但是对应于第一存储装置内设备的设备称为虚拟设备。此时，将第三设备作为虚拟设备，从第一存储装置内的第一设备转移到第二存储装置内的第二设备。

作为一个实施方式，定义在第二存储装置中将第三设备虚拟化的第二设备，通过交接第一设备的属性等、与主机联动，将成为访问对象的设备从第一设备切换到第二设备，由此将虚拟化控制存储装置中的虚拟设备的定义(以下，称为虚拟设备定义)从第一存储装置转移到第二存储装置。

另外，作为别的实施方式，定义在第二存储装置中将第一设备虚拟化的第二设备，将来自主机的访问对象从第一设备切换到第二设备后，将第二设备的虚拟化对象从第一设备变更为第三设备，由此进行虚拟设备定义的转移。

再有，为了重新调整存储装置处理能力和输入输出负载的平衡，管理服务器周期性地收集第一以及第二存储装置的构成信息或动作信息，监视第一以及第二存储装置。在检测基于控制处理器或磁盘超高速缓存存储器等的构成部位故障的 I/O 处理能力的低下、或基于存储装置的新追加或机器老朽化的撤消等系统构成变更、或基于来自主机的 I/O 负载变动的主机负载的偏移的发生的基础上，进行上述的设备的转移，由此实现将存储装置之间的 I/O 负载均匀化的系统。

根据本发明，可以继续运用存储装置系统的同时，进行虚拟化控制存储装置的替换或增、减设。

再有，根据主机的 I/O 负载变动等，能够将虚拟化控制存储装置之间的输入输出负载最佳化。

#### 附图说明

图 1 是表示应用本发明的计算机系统的硬件结构的一例的图；

图 2 是表示应用本发明的计算机系统的软件结构的一例的图；

图 3 是表示逻辑设备管理信息的一例的图；

图 4 是表示 LU 路径管理信息的一例的图；

图 5 是表示物理设备管理信息的一例的图；

图 6 是表示外部设备管理信息的一例的图；

图 7 是表示在本发明实施例 1 中、在管理服务器 110 执行的设备移动指示处理 241 的处理流程的一例的图；

图 8 是表示在本发明实施例 1 中、在存储装置 130 执行的设备移动处理 222 的处理流程的一例的图；

图 9 是表示在本发明实施例 1 中、在存储装置 130 执行的存储装置 I/O 处理 221 的处理流程的一例的图；

图 10 是表示在本发明实施例 1 中、在主机 100 执行的主机 I/O 处理 261 的处理流程的一例的图；

图 11 是表示在本发明实施例 1 或 2 中、在管理服务器 110 执行的存储装置负载监视处理 242 的处理流程的一例的图；

图 12 是表示在本发明实施例 2 中、在管理服务器 110 执行的设备移动指示处理 241 的处理流程的一例的图；

图 13 是表示在本发明实施例 2 中、在存储装置 130 执行的设备移动处理 222 的处理流程的一例的图；

图 14 是表示在本发明实施例 2 中、在存储装置 130 执行的存储装置 I/O 处理 221 的处理流程的一例的图；

图 15 是表示应用本发明的计算机系统的硬件结构的另外一例的图；

图 16 是表示在本发明实施例的变形例中、在管理服务器 110 执行的设备移动指示处理 241 的处理流程的一例的图。

#### 具体实施方式

作为本发明的实施例，说明第一以及第二实施例。

第一以及第二实施例是，在基于将第三存储装置的设备作为虚拟设备虚拟化的多个存储装置的系统（以下，称为虚拟存储装置系统）中，实施使用了外部存储装置连接技术的第一以及第二存储装置的例子。

第一实施例中，在虚拟存储装置系统中，在将第三存储装置的第三设备作为第一存储装置的第一设备来虚拟化时，在第二存储装置中定义将第三设备虚拟化的第二设备，将第一设备的属性等从第一存储装置转移到第二存储装置。然后，通过和主机联动，将来自主机的称为访问对象的设备从第一设

备切换到第二设备，由此进行虚拟设备定义的转移。

具体地说，首先，管理服务器根据来自用户或存储装置管理者的指示，在成为移动目的地的第二存储装置中，定义将第三设备虚拟化的第二设备，将设定在第一设备的各种属性，设定在第二设备中，进行第二设备的 LU 路径定义。这里，所谓 LU 路径定义，是在各端口的 LUN (Logical Unit Number) 对应逻辑设备，以使上级装置等能够从存储装置的端口访问逻辑设备。然后，服务器装置对主机进行使主机识别第二设备的指示后，将对第一存储装置的第一设备的输入输出处理动作模式设定为和主机请求同步地执行向磁盘装置的写入的模式（以下，称为超高速缓冲存储器直通模式）。

接着，管理服务器对主机，将向第二设备的路径定义为向第一设备的路径的交替路径。这里，所谓交替路径，是从主机对于具有多个路径（路径）的逻辑设备，指示从多个路径中用于访问的路径切换到其他路径来访问时使用的其他路径。然后，对第一存储装置指示向第一设备的 I/O 停止。结果，主机将其后的向第一设备的 I/O 通过作为交替路径设定的路径转换到第二设备。然后，管理服务器向主机请求第一设备的交替路径定义的解除，解除第一存储装置的第一设备的定义。

第二实施例中，在上述的虚拟存储装置系统中，在将第三存储装置的第三设备作为第一存储装置的第一设备虚拟化时，在第二存储装置中定义将第一设备虚拟化的第二设备，将主机的访问对象从第一设备切换到第二设备后，将第二设备的虚拟化对象从第一设备变更为第三设备，由此进行虚拟设备定义的转移。

具体地说，首先，服务器装置根据来自用户或者存储装置管理者的指示，在成为移动目的地的第二存储装置中定义将第一设备虚拟化的第二设备，将设定在第一设备的各种属性，设定在第二设备，进行第二设备的 LU 路径定义。然后，进行使主机识别第二设备的指示后，对第二存储装置，将第二设备的输入输出处理动作模式设定为超高速缓冲存储器直通模式。

接着，对主机，将向第二设备的访问路径定义为第一设备的交替路径。然后，对第一存储装置指示来自主机的向第一设备的 I/O 停止。结果，主机将其后的向第一设备的 I/O 经由设定为交替路径的路径调换为第二设备。然

后，解除主机的第一设备的交替路径的定义和第一存储装置的第一设备的定义。再有，对第二存储装置指示解除第二设备的超高速缓存存储器直通模式，对第一存储装置指示设定第一设备的超高速缓存存储器直通模式。然后变更设备的对应关系，以使第二存储装置的第二设备将第三存储装置的第三设备虚拟化。

在第一实施例中，在将第三设备虚拟化的第一以及第二存储装置两方进行超高速缓冲存储器直通模式的动作，因为同步进行其间的对第三存储装置的第三设备的写入处理，因此引起性能低下。但是，在第二实施例中，一旦，在第二存储装置中将第一存储装置的虚拟设备虚拟化（即，双重虚拟化），因此在转移处理中，在第一或第二存储装置的至少一方可进行使用了磁盘超高速缓冲存储器的异步写入处理，可以减轻性能低下。

再有，在第一以及第二实施例中，管理服务器 110 监视第一以及第二存储装置的动作信息或构成信息，做成检测负载向特定的虚拟设备或第一以及第二存储装置集中的结构，由此在第一以及第二存储装置之间能够实现输入输出负载的均匀化。

### 实施例 1

首先，参照图 1 至图 10 说明第一实施例。

图 1 是表示适用第一实施例的计算机系统的硬件结构的一例的图。

计算机系统包括：一台或其以上的主机 100、管理服务器 110、光纤信道开关 120、存储装置 130a 和 130b（统称为存储装置 130）、管理终端 140a 和 140b（统称为管理终端 140）以及外部存储装置 150a 和 150b（统称为外部存储装置 150）。主机 100、存储装置 130 以及外部存储装置 150，通过各个端口 107、131、151 连接到光纤信道开关 120 的端口 121。另外，主机 100、存储装置 130、外部存储装置 150 和光纤信道开关 120，从各个接口控制部（I/F）106、138、157、123，通过 IP 网络 175 连接到管理服务器 110，通过在管理服务器 110 动作的图中未示的存储装置管理软件进行管理。此外，在本实施例中，假定采取存储装置 130 通过管理终端 140 连接到管理服务器 110 的形态，但也可以是存储装置 130 被直接连接到 IP 网络的结构。

主机 100 是具有 CPU101 或内存 102 等的计算机。主机 100 将存储在磁

盘装置或光磁盘装置等存储装置 103 的操作系统或应用程序等软件读入到内存 102, CPU101 从内存 102 读出后执行, 由此执行规定的处理。主机 100 具备键盘或鼠标等输入装置 104 或显示器 105 等输出装置, 通过输入输出装置 104 接受来自主机管理者等的输入, 通过输出装置输出由 CPU101 指示的信息。另外, 主机 100 除了端口 107, 还具有用来和 IP 网络 175 连接的一个或其以上的接口控制部 106。

管理服务器 110 和主机 100 相同, 是具有 CPU111 或内存 112 的计算机。管理服务器 110, 将存储在磁盘装置或光磁盘装置等存储装置 113 的存储装置管理软件等读入到内存 112, CPU111 从内存 112 读出其软件后执行, 由此执行诸如计算机系统整体运用/维修管理的规定的处理。由 CPU111 执行存储装置管理软件后, 管理服务器 110 从接口控制部 116 通过 IP 网络 175 收集来自计算机系统内各机器的构成信息、资源利用率、性能监视信息和故障日志等。然后, 管理服务器 110, 将收集到的这些信息输出到显示器 115 等输出装置, 提示给存储装置管理者。另外, 管理服务器通过键盘或鼠标等输入装置 114 接收来自存储装置管理者的指示, 通过接口控制部 116 将接收到的运用/维修指示发送到各机器。

存储装置 130 具有: 一个或多个端口 131、一个或多个控制处理器 132、连接到各个控制处理器 132 的一个或多个内存 133、一个或多个磁盘超高速缓冲存储器 134、一个或多个控制内存 135、一个或多个端口 136、连接到各个端口 136 的一台或其以上的磁盘装置 137 和接口控制部 138。

控制处理器 132 对从端口 131 接收到的输入输出请求指定访问对象设备, 处理向对应于该设备的磁盘装置 137 或外部存储装置 150 内的设备的输入输出请求。此时, 控制处理器 132 根据接收到输入输出请求中包含的端口 ID 以及 LUN (Logical Unit Number), 指定访问对象设备。此外, 在本实施例中, 作为端口 131, 假定对应于将 SCSI (Small Computer System Interface) 作为上级协议的光纤信道接口的端口。但是, 也可以是对应于将 SCSI 作为上级协议的 IP 网络接口等网络接口的端口。

本实施例的存储装置 130, 具有如下所述的设备阶层。首先, 构成基于多台磁盘装置 137 的磁盘阵列, 该磁盘阵列被控制处理器 132 作为物理设备

管理。再有，控制处理器 132 对装载在存储装置 130 内的物理设备分配逻辑设备（即，控制处理器 132 将物理设备和逻辑设备对应起来）。在存储装置 130 内管理逻辑设备，其号码对于每一存储装置 130 独立进行管理。

逻辑设备对应于分配给各端口 131 的 LUN，将逻辑设备作为存储装置 130 的设备提供给主机 100。即，主机识别的是存储装置 130 的逻辑设备，主机 100 利用与逻辑设备对应的端口 131 的 LUN，访问存储在存储装置 130 的数据。此外，在本实施例中，控制处理器 132 将外部存储装置 150 的逻辑设备作为外部设备（以下，对于某存储装置将处于另外筐体的存储装置内的逻辑设备叫做该存储装置的外部逻辑设备）管理，作为存储装置 130 的设备来虚拟化。即，由存储装置 130 管理的外部设备是虚拟设备。利用外部存储装置连接技术引进的外部设备和物理设备一样，一个或其以上的外部设备对应于该存储装置 130 的逻辑设备。此外，外部设备也在各存储装置 130 内被独立管理。

为了实现这样的设备阶层，控制处理器 132 对逻辑设备、物理设备、磁盘装置 137、外部设备、外部存储装置 150 的逻辑设备进行管理，并且还管理各个设备之间的对应关系。而且，控制处理器 132 将对逻辑设备的访问请求变换为向磁盘装置 137 或外部存储装置装置 150 的逻辑设备的访问请求，进行向适当的装置发送的处理。

此外，和上述的说明一样，本实施例中的存储装置 130，将多个磁盘装置 137 归纳起来定义为一个或多个物理设备（即，将多个磁盘装置 137 归纳起来，使其与一个或多个物理设备对应），对一个物理设备分配一个逻辑设备，提供给主机 100。但是，也可以做成让主机 100 将各个磁盘装置 137 作为一个物理设备以及一个逻辑设备来看待。

另外，控制处理器 132 除了对设备的输入输出外，还执行数据复制或数据再配置等设备之间的各种数据处理。

另外，控制处理器 132 向通过接口控制部 138 连接的管理终端 140 发送向存储装置管理者提示的构成信息，由管理终端 140 接收管理者向管理终端 140 输入的维修/运用指示后进行存储装置 130 的构成变更等。

为了提高从主机 100 对数据的访问请求的处理速度，在磁盘超高速缓冲

存储器 134 中预先存储从磁盘装置 137 频繁地读出的数据、或临时地存储从主机 100 接收到的数据。此外，在磁盘超高速缓冲存储器 134 中存储从主机 100 接收到的写入用数据之后，在实际写入到磁盘装置 137 之前，对主机 100 返回对写入请求的应答的处理叫写后处理（write-after）。

在进行写后处理（write-after）的情况下，理想的是，为了防止存储在磁盘超高速缓冲存储装置中的数据写入到磁盘装置 137 之前消失，通过电池备份等使磁盘超高速缓冲存储器 134 不挥发，或者为了提高对介质故障的抵抗性，配置两个磁盘超高速缓冲存储器 134，提高磁盘超高速缓冲存储器 134 的可用性。

控制内存 135 中，存储用来实现上述的设备阶层的各设备的属性、用来管理设备之间的对应关系的控制信息、用来管理在磁盘超高速缓冲存储器 134 上保存的磁盘反映完毕或未反映数据的控制信息等。

若存储在控制内存 135 中的控制信息消失，则不能访问存储在磁盘装置 137 的数据，因此，理想的是，通过电池备份等使控制内存 135 不挥发，或者为了提高对介质故障的抵抗性，配置两个控制内存 135，提高可用性。

存储装置 130 内的各部位如图 1 所示，用内部结合网连接，执行各部位之间的数据、控制信息以及构成信息的收发。通过该内部结合网，控制处理器 132 之间能够共享存储装置 130 的构成信息来管理。此外，从可用性提高的观点来看，理想的是配置多个内部结合网。

管理终端 140，具有 CPU142、内存 143、存储装置 144、和存储装置 130 连接的接口控制部 141、和 IP 网络 175 连接的接口控制部 147、接收来自存储装置管理者的输入的输入装置 145、以及向存储装置管理者输出存储装置 130 的构成信息或管理信息的显示器 146 等输出装置。CPU142 将存储在存储装置 144 的存储装置管理程序读出到内存 143 后执行，由此，进行构成信息的参照、构成变更的指示、特定功能的动作指示等，就其存储装置 130 的维修运用而言，成为存储装置管理者或管理服务器 110 和存储装置 130 之间的接口。此外，在本实施方式中，也可以不经由管理终端 140 而将存储装置 130 直接连接到管理服务器 110，使用在管理服务器 110 动作的管理软件来管理存储装置 130。

外部存储装置 150，具有：通过光纤信道开关 120 和存储装置 130 的端口 131 连接的端口 151、控制处理器 152、内存 153、磁盘超高速缓冲存储器 154、一个或多个磁盘装置 156 以及和各个磁盘装置连接的一个或多个端口 155。控制处理器 152 执行存储在内存 153 的程序，由此处理从端口 151 接收到的、向磁盘装置 156 的输入输出请求。此外，在本实施方式中，外部存储装置 150，做成为不具有控制内存，且比存储装置 130 小的规模的结构的存储装置，但是也可以是具有和存储装置 130 一样的结构的一个或多个相同规模的存储装置。

此外，在本实施例中，如图 1 所示，因为通过光纤信道开关 120 连接存储装置装置 130 的端口 131 和外部存储装置 150 的端口 151，因此最好设定光纤信道开关 120 的分区，以使抑制来自主机 100 的对外部存储装置 150 的直接访问。另外，也可以不经由光纤信道开关 120 而直接连接端口 131 和端口 151。

接着，对实施例中的存储装置 130 的软件结构进行说明。图 2 是表示存储在存储装置 130、管理服务器 110、主机 100 的控制内存以及内存中的控制信息和存储装置控制处理用的程序的一例的软件结构图。

作为存储装置 130 的构成管理信息，控制内存 135 中有逻辑设备管理信息 201、LU 路径管理信息 202、物理设备管理信息 203、外部设备管理信息 204、超高速缓冲存储器管理信息 205、设备功能管理信息 206。在本实施例中，为了防止信息消失，将这些信息存储在控制内存 135 中，可以由控制处理器 132 参照、更新，但是此时需要经过相互连接网的访问。

在本实施例中，为了提高处理性能，将在各控制处理器 132 中执行的处理中所需要的控制信息的复制保存到内存 133。另外，管理终端 140 或管理服务器 110 也具有存储装置 130 的构成管理信息的复制。在从管理服务器 110 或管理终端 140 接收来自存储装置管理软件或存储装置管理者的指示后变更了存储装置 130 的结构的情况下、或存储装置 130 内各部位通过故障/自动交替等而结构发生了变化的情况下，控制处理器 132 中的一个，更新控制内存 135 内的相应构成管理信息。然后，控制处理器 132 将控制信息通过结构变更被更新的旨意通过相互连接网通知给其他控制处理器 132、管理终端 140、

管理服务器 110，从控制内存 135 向其他部位的内存读入最新信息。

另外，管理服务器 110 将从存储装置 130 或外部存储装置 150 接收到的设备管理信息的复制 231 或表示各存储装置的属性的、图中未示存储装置管理信息保存在内存 112。为了避免数据的消失，也可以将这些信息保存在装载在管理服务器 110 的存储装置 113 中。

主机 100 识别存储装置 130 或外部存储装置 150 提供的逻辑设备，作为设备文件来管理。将设备文件和各存储装置的逻辑设备之间的对应管理信息叫做设备路径管理信息 251。在存储装置 130 从多个端口 131 提供同一个逻辑设备情况下，对同一个逻辑设备生成多个设备文件。在主机 100 中，根据路径管理软件等，统一管理对同一个逻辑设备的多个设备文件，进行向该逻辑设备的输入输出负载分散或端口故障时的交替处理。

接着，对各管理信息进行说明。第一，对逻辑设备管理信息 201 进行说明。图 3 表示逻辑设备管理信息 201 的一例。逻辑设备管理信息 201，关于各逻辑设备，保存从记录 301 到记录 311 的信息组。

记录 301 中存储为了识别逻辑设备、控制处理器 132 分配给逻辑设备的号码（以下，称为逻辑设备号码）。

记录 302 中作为尺寸存储有根据逻辑设备号码指定的逻辑设备的容量。

记录 303 中存储对应于该逻辑设备的物理设备或外部设备的号码。即，存储作为各个管理信息的物理设备管理信息 203 或外部设备管理信息 204 的记录 501 或记录 601。在未定义逻辑设备的情况下，在记录 303 中设定无效值。

此外，在本实施例中，因为逻辑设备和物理/外部设备是一对一对应的，因此记录 303 中存储一个物理/外部设备的号码。但是，在结合多个物理/外部设备来形成一个逻辑设备的情况下，必须有存储各逻辑设备对应的物理/外部设备的号码一览表和其数目的记录。此外，未定义逻辑设备的情况下，记录 303 中设定无效值。

记录 304 中存储该逻辑设备的设备类型识别信息。存储装置 130 能够定义在超高速缓冲存储器上的数据管理单位或设备管理信息的存储形态（有没有向磁盘空间存储管理信息、存储形态等）不同的多个设备类型的逻辑设备。

因此，在记录 304 中记录各逻辑设备是哪一种设备类型。

记录 305 中设定表示该逻辑设备的状态的信息（以下，称为设备状态）。作为设备状态，有“在线”、“离线”、“未安装”、“已封锁”。“在线”表示该逻辑设备正常工作、通过管理服务器 110 在一个或其以上的端口 131 定义 LU 路径、能够从主机 100 访问的状态。“离线”表示定义了该逻辑设备并且正常工作、但是由于未定义 LU 路径、因此不能从主机 100 访问的状态。“未安装”表示对物理设备或外部设备未定义该逻辑设备、不能从主机 100 访问的状态。“已封锁”表示该逻辑设备发生了故障而不能从主机 100 访问的状态。记录 305 的初始值为“未安装”，通过逻辑设备定义处理变更为“离线”，进而通过 LU 路径定义处理变更为“在线”。

记录 306 中存储端口号码、目标 ID、LUN 一览表等识别该逻辑设备的信息。端口号码是表示对应于该逻辑设备的 LUN 是多个端口 131 中的哪一个端口的信息，即，是为了访问该逻辑设备而使用的端口 131 的识别信息。这里，所谓端口 131 的识别信息，是分配给各端口 131 的、唯一在存储装置 130 内能识别的号码。另外，存储在记录 306 的目标 ID 和 LUN 一览表是用来识别该逻辑设备的标识符。在本实施例中，作为用来识别逻辑设备的标识符，使用在 SCSI 上从主机 100 访问设备的情况下使用的 SCSI-ID 和 LUN。记录 306 中的各信息是在执行对该逻辑设备的定义时被设定。

记录 307 中存储识别允许向该逻辑设备的访问的主机 100 的主机名（以下，称为连接主机名）。作为主机名，只要是对主机 100 的端口 107 赋予的 WWN（World Wide Name）等唯一能够识别主机 100 或端口 107 的值都可以使用。除了这些，存储装置 130 还具有各端口 131 的 WWN 等关于端口属性的管理信息。该记录 307 是由存储装置管理者定义逻辑设备时被设定。

记录 308 中，在该逻辑设备在后述的同步复制中或转移中的情况下，设定复制目的地或转移目的地的物理设备号码或外部设备号码，在该逻辑设备既不在同步复制中、也不在数据转移中的情况下，设定无效值。

记录 309 中，在该逻辑设备在数据转移中的情况下，存储表示数据转移结束的最终地址的信息（以下，称为数据转移进展指针）。数据转移进展指针在接受从主机 100 向该逻辑设备的 I/O 请求时，判定访问对象区域是转移完

成还是未完成，在选择处理时使用。

记录 310 中存储表示该逻辑设备在同步复制中还是数据转移中还是两种状态都不对应的信息（以下，称为同步复制/数据转移中标志）。

记录 311 中存储表示是向该逻辑设备的数据更新处理否将通过超高速缓冲存储器直通模式进行的信息（以下，称为超高速缓冲存储器直通标志）。所谓超高速缓冲存储器直通模式，具体地说，表示在有来自主机的数据更新时将更新数据存储在超高速缓冲存储器中、进而向磁盘装置写入该更新数据、其后将该更新处理的结束向主机报告的动作。

第二，对 LU 路径管理信息 202 进行说明。图 4 表示 LU 路径管理信息 202 的一例。LU 路径管理信息 202 关于存储装置 130 内的各端口 131，保存在各端口定义的有效 LUN 数量的信息。

记录 401 中如上述的端口号那样存储识别该端口 131 的信息。

记录 402 中存储在端口 131 定义的（分配的）LUN。记录 403 中存储该 LUN 被分配的逻辑设备的号码（以下，称为对应逻辑设备号码）。记录 404 中存储表示允许访问在该端口 131 定义的该 LUN 的主机 100 的信息。作为表示主机 100 的信息，例如使用对上述的主机 100 的端口 107 赋予的 WWN。

此外，有时对一个逻辑设备定义多个端口 131 的 LUN，从多个端口 131 能够访问该逻辑设备。这种情况下，关于该多个端口 131 的各个 LUN 的 LU 路径管理信息 202 的连接主机名的并集（一览表），被保存在关于该逻辑设备的逻辑设备管理信息 201 的记录 307 中。

第三，对物理设备管理信息 203 进行说明。物理设备管理信息 203 用于由存储装置 130 内的一台或其以上的磁盘装置 137 构成的物理设备的管理。图 5 表示物理设备管理信息 203 的一例。各存储装置 130 对位于自身装置内的每个物理设备，保存从记录 501 到记录 507 的信息组。

记录 501 中登录用来识别物理设备的识别号码（以下，称为物理设备号码）。记录 502 中存储由记录 501 指定的物理设备的容量。记录 503 中将该物理设备对应的逻辑设备号码作为对应逻辑设备号码存储。记录 503 的存储是在定义该物理设备对应的该逻辑设备时进行。在该物理设备未被分配给逻辑设备的情况下，记录 503 中设定无效值。

记录 504 中设定表示该物理设备的状态的信息（设备状态）。作为状态，有“在线”、“离线”、“未安装”、“已封锁”。“在线”表示该物理设备正常工作并且被分配给逻辑设备的状态。“离线”表示定义该物理设备且该物理设备正常工作、但是未分配给逻辑设备。“未安装”表示在磁盘装置 137 上没有定义该物理设备的状态。“已封锁”表示该物理设备发生了故障而不能访问的状态。记录 504 的初始值为“未安装”，通过物理设备定义处理变更为“离线”，进而以定义逻辑设备为契机变更为“在线”。

记录 505 中保存有分配了该物理设备的磁盘装置 137 的 RAID 级、数据磁盘和奇偶校验磁盘数、作为数据分割单位的磁条的尺寸等与 RAID 结构有关的信息。记录 506 中将构成分配了该物理设备的 RAID 的多个磁盘装置 137 的各自的识别号码做成一览表保存。该磁盘装置 137 的识别号码是为了在存储装置 130 内识别磁盘装置 137 而赋予的唯一的值。记录 507 中存储表示该物理设备是被分配给各磁盘装置 137 内的哪一个区域的信息。在本实施例中，为了简便起见，对于全部物理设备，设和构成 RAID 的各磁盘装置 137 内的偏移相同的尺寸。

第四，对外部设备管理信息 204 进行说明。外部设备管理信息 204 是为了将连接到存储装置 130 的外部存储装置 150 的逻辑设备作为外部设备管理而使用的。图 6 表示外部设备管理信息 204 的一例。存储装置 130 对位于外部存储装置 150 内的每一个逻辑设备（=外部设备）保存从记录 601 到记录 608 的信息组。

记录 601 中存储存储装置 130 的控制处理器 132 对该外部设备分配的、在存储装置 130 内唯一的值（以下，称为外部设备号码）。记录 602 中存储由外部设备号码指定的外部设备的容量。记录 603 中将该外部设备对应的存储装置 130 内的逻辑设备的号码作为对应逻辑设备号码来登录。

记录 604 中，和物理设备管理信息 203 的记录 504 一样，设定该外部设备的设备状态。此外，在初始状态中，因为存储装置 130 没有连接外部存储装置 150，因此记录 604 的初始值为“未安装”。

记录 605 中存储装载该外部设备的外部存储装置 150 的识别信息（以下，称为存储装置识别信息）。作为存储装置识别信息，考虑该存储装置的供应商

识别信息和各供应商唯一分配的制造系列号码的组合等等。记录 606 中，对于与该外部设备对应的外部存储装置 150 内的逻辑设备，在外部存储装置 150 内分配的识别号码、即逻辑设备号码作为外部存储装置内设备号码被存储。

记录 607 中登录可访问该外部设备的存储装置 130 的端口 131 的识别号码（以下，称为启动程序端口号一览表）。从多个端口 131 能够访问该外部设备的情况下，登录多个端口识别号码。

记录 608 中，在该外部设备在外部存储装置 150 的一个或其以上的端口 151 定义 LUN 的情况下，保存这些端口 151 的端口 ID 以及分配了该外部设备的一个或多个目标 ID/LUN。此外，在存储装置装置 130 的控制处理器 132 通过端口 131 访问外部设备的情况下，作为用来识别该外部设备的信息，使用由该外部设备所属的外部存储装置 150 分配给该外部设备的目标 ID 以及 LUN。

第五，对设备功能管理信息 206 进行说明。本信息中存储表示对各逻辑设备设定的各种属性的信息。作为设备的属性信息的例子，可以列举将向该逻辑设备的访问只限定在特定端口的访问控制信息、抑制向该逻辑设备的读写访问的访问属性信息、向该逻辑设备内数据的有无适用加密或加密/解密中使用的密钥信息等加密设定信息等等。

在本实施例中，存储装置 130 使用上述的各设备管理信息来管理设备。作为该存储装置 130 的初始状态，假定该装置出厂时已经在各设备装置 137 定义了物理设备。用户或存储装置管理者在导入存储装置 130 时，将在该存储装置 130 上连接的外部存储装置 150 的逻辑设备作为外部设备来定义。另外，对这些物理设备或者外部设备定义逻辑设备，就其该逻辑设备来说是在各端口 131 定义 LUN。

接着，再次回到图 2，对存储在存储装置 130、管理服务器 110、主机 100 的内存 133、112、102 中存储的程序进行说明。在本实施例中，按照来自用户或存储装置管理者的指示，在存储装置 130 间进行切换外部设备的控制担当的处理和对该切换处理中的逻辑设备的主机 I/O 处理。这些处理是管理服务器 110、存储装置 130、主机 100 联合起来执行的。

首先，对应外部设备的控制担当切换处理，在管理服务器 110 中，设备

移动指示处理 241 存储在内存 112；在存储装置 130 中，设备移动处理 222 存储在内存 133。另外，对应向切换中的逻辑设备的主机 I/O 处理，在主机 100 中，主机 I/O 处理 261 存储在内存 102；在存储装置 130 中存储装置 I/O 处理 221 存储在内存 133。

图 7 表示在管理服务器 110 执行的、设备移动指示处理 241 的处理流程的一例的图。管理服务器 110 通过 IP 网络 175 和接口控制部 116 从用户或管理者或应用程序受理将存储装置 130a（以下，称为移动源存储装置）的第一逻辑设备移动至存储装置 130b（以下，称为移动目的地存储装置）的第二逻辑设备的请求。作为包含在该请求中的信息，例如考虑存储装置 130a 以及 130b 的识别信息、第一以及第二逻辑设备号码、存储装置 130b 中的第二逻辑设备的 LU 路径定义中所需要的信息（存储装置 130b 的端口 131 特定信息以及 LUN 一览表）等等。

管理服务器 110，判定与移动对象的逻辑设备对应的设备是外部设备或该存储装置 130 的内部设备（物理设备）中的某一个（步骤 701）。

管理服务器 110，在该逻辑设备对应于外部设备的情况下，对外部存储装置 150 指示第三逻辑设备的访问权变更，以使从成为移动目的地的存储装置 130b 能够访问与该外部设备对应的外部存储装置 150 的第三逻辑设备（步骤 702）。

管理服务器 110 在移动目的地存储装置 130b 定义将第三逻辑设备虚拟化的外部设备，对应于该外部设备定义第二逻辑设备（步骤 703）。作为外部设备定义处理，具体地说，关于在存储装置 130b 将第三逻辑设备虚拟化的外部设备，在对应的外部设备管理信息 204 的记录中，在记录 602 到记录 608 设定关于外部存储装置 150 以及第三逻辑设备的信息。此时，管理服务器 110 将记录 604 的设备状态从“未安装”变更为“离线”。另外，管理服务器 110 对于与存储装置 130b 的第二逻辑设备对应的逻辑设备管理信息 201 的各记录设定控制信息。例如，对于记录 303 的对应物理/外部设备号码，设定将第三逻辑设备虚拟化的该外部设备的号码。

另外，如上所述，在本步骤 703 中，将外部设备管理信息 204 的记录 604 变更为“在线”，将逻辑设备管理信息 201 的记录 305 从“未安装”变更为“离

线”。另外，在第二逻辑设备的逻辑设备管理信息 201 中，记录 311 的超高速缓冲存储器直通标志被设定为接通，在移动目的地存储装置 130b 中，以超高速缓冲存储器直通模式处理第二管理设备的写入处理。

管理服务器 110，将在移动源存储装置 130a 的第一逻辑设备设定的各种属性，设定在移动目的地存储装置 130b 的第二逻辑设备（步骤 704）。此时，作为从第一逻辑设备向第二逻辑设备交接的各种属性的例子，可以列举将向该逻辑设备的访问只限定在特定主机 100 的访问控制信息、抑制向该逻辑设备的读写访问的访问属性信息和访问属性信息的设定日期和有效期间等附带信息、用于向该逻辑设备内数据的有无适用加密和加密/解码中的密钥信息等加密设定信息等等。

在本步骤 704 中，管理服务器 110 从移动源存储装置 130a 取得有关第一逻辑设备的属性，向移动目的地存储装置 130b 的第二逻辑设备设定相同的属性。例如，在第一逻辑设备设定只允许来自特定的主机 100 的访问的访问属性的情况下，管理服务器取得第一逻辑设备的逻辑设备管理信息 201 中的记录 307 的连接主机名的值，向第二逻辑设备的逻辑设备管理信息 201 中的记录 307 设定。另外，在第一逻辑设备设定在 2010 年 1 月 1 日禁止 10 年期间的写入访问的访问属性的情况下，管理服务器 110 对第二逻辑设备设定从 2010 年 1 月 1 日设定的日期开始禁止 10 年期间的写入访问的访问属性。

另外，在本步骤 704 中，将第一逻辑设备的属性向第二逻辑设备设定的处理并不是由管理服务器 110 来进行，而可以在移动源存储装置 130a 和移动目的地存储装置 130b 之间交接设备属性。例如，可以从移动源存储装置 130a 对移动目的地存储装置 130b 询问第一逻辑设备的属性信息，将从移动源存储装置 130a 返回的设备属性信息设定在移动目的地存储装置 130b 的第二逻辑设备。另外，相反，也可以从移动源存储装置 130a 向移动目的地存储装置 130b 转发第一逻辑设备的属性，指示向第二逻辑设备的属性设定。

然后，管理服务器 110 在移动目的地存储装置 130b 定义第二逻辑设备的 LU 路径（步骤 705）。具体地说，对于与 LU 路径定义对象的端口 131 对应的 LU 路径管理信息 202 的记录，进行记录 403 的对应逻辑设备号码等的设定。另外，在本步骤 705，在与第二逻辑设备对应的逻辑设备管理信息 201 的记

录，设定定义了 LU 路径的端口 131 信息，将记录 305 变更为“在线”。

管理服务器 110 对主机 100 进行使它识别与第二逻辑设备对应的移动目的地存储装置 130b 的 LU 的指示（步骤 706）。这里，接收到来自管理服务器 110 的 LU 识别指示的主机 100，对于与管理服务器 110 新分配给存储装置 130b 的第二逻辑设备对应的 LU，制作设备文件。例如，在 Hewlett Packard 公司的 UNIX（登录商标）操作系统中，通过“IOSCAN”命令，进行新 LU 的识别和设备文件的制作。

在主机 100，若对应于新制作的设备文件的存储装置 130 的逻辑设备和已制作的设备文件相同，则对它进行检测，将这些设备文件作为同一个组来管理。在该同一性的判定中考虑通过 SCSI 的 Inquiry 命令等取得与 LU 对应的各存储装置 130 内的逻辑设备号码的方法。但是，第一逻辑设备和第二逻辑设备是分别不同的存储装置 130a 和 130b 的设备，因此在该时刻不能作为一个组来管理，主机 100 将两个设备文件作为不同的组来识别。在本步骤 706 中，管理服务器 110，在一台或其以上的主机 100 访问与移动源存储装置 130a 的第一逻辑设备对应的 LU 的情况下，对全部主机 100，进行同样地识别移动目的地存储装置 130b 的 LU 的指示。

管理服务器 110，如果由主机 100 进行的移动目的地存储装置 130b 的 LU 识别结束了，则指示移动源存储装置 130a 将第一逻辑设备的动作模式变更为超高速缓冲存储器直通模式（步骤 707）。通过该动作模式变更，存储装置 130a，设备移动处理 222 将关于第一逻辑设备、在磁盘超高速缓冲存储器 134 中保存的未反映的更新数据写入到全部外部存储装置 150 的第三逻辑设备。通过本步骤 707，来自主机 100 的向第一逻辑设备的数据更新，与来自主机的更新请求同步地向外部存储装置 150 的第三逻辑设备写入。因此，能够避免未写入的更新数据停留在存储装置 130a 内磁盘超高速缓冲存储器 134 内。关于在存储装置 130 执行的、在超高速缓冲存储器直通模式中的更新数据的写入处理，在后述的设备移动处理 222 中进行说明。

管理服务器 110 若从移动源存储装置 130a 接收到伴随着向超高速缓冲存储器直通模式的变更的上述写入处理结束的通知，则对主机 100 进行将对与第二逻辑设备对应的移动目的地存储装置 130b 的 LU 的路径定义为关于移动

源存储装置 130a 的第一逻辑设备的交替路径的指示（步骤 708）。此时，一台或其以上的主机 100 访问移动源存储装置 130a 的第一逻辑设备的情况下，对这些所有的主机 100 也同样进行指示，以定义交替路径。通过该交替路径定义，主机 100 将在步骤 706 制作的两个设备文件作为对应于同一个逻辑设备的文件来识别。这样，若作为一个交替路径的组来定义，则主机 100 从经由存储装置 130a 以及 130b 的两个 LU 路径可以访问外部存储装置 150 的第三逻辑设备。

交替路径定义结束后，管理服务器 110 指示移动源存储装置 130a 拒绝接受向第一逻辑设备的来自主机 100 的 I/O 请求（步骤 709），进而对主机 100 进行解除移动源存储装置 130a 的 LU 交替路径定义的指示（步骤 710）。此时，从一台或其以上的主机 100 访问第一逻辑设备的情况下，对这些所有的主机 100 进行同样的指示。

管理服务器 110 指示存储装置 130b 解除第二逻辑设备的超高速缓冲存储器直通模式，允许使用了磁盘超高速缓冲存储器的写入处理（步骤 711）。

管理服务器 110 进行来自移动源存储装置 130a 的、解除将外部存储装置 150 的第三逻辑设备虚拟化的外部设备的定义的指示（步骤 712），对外部存储装置 150 变更第三逻辑设备的访问权，以禁止来自移动源存储装置 130a 的、对第三逻辑设备的访问（步骤 713）。

另一方面，对在步骤 701 的处理中对象逻辑设备对应于存储装置 130a 内部的物理设备的情况进行说明。

管理服务器 110，首先从移动目的地存储装置 130b 的物理设备中没有定义逻辑设备的未使用的物理设备中，选定成为移动目的地的设备，对应本物理设备定义第二逻辑设备（步骤 714）。此时，也可以作为设备移动请求的参数信息，接收在移动目的地存储装置 130b 的移动目的地物理设备特定信息，作为移动目的地采用被指示的物理设备。另外，在本实施例中，作为移动目的地设备选定移动目的地存储装置 130b 内的物理设备，但是也可以将移动目的地存储装置 130b 所管理的外部设备作为移动目的地设备来选定。

接着，管理服务器 110 在第二逻辑设备设定作为移动源的第一逻辑设备的属性（步骤 715），在移动目的地存储装置 130b 指示第二逻辑设备的 LU 路

径定义（步骤 716）。

如果使一台或其以上的主机 100 识别该 LU（步骤 717），则指示移动源存储装置 130a，与移动目的地存储装置 130b 的第二逻辑设备同步复制第一逻辑设备的所有数据（步骤 718）。所谓同步复制，是指与主机的写入 I/O 请求同步地进行复制处理。在同步复制中，来自主机的写入处理结束时，必须保证复制处理也结束。接受本指示，移动源存储装置 130a 执行对象逻辑设备的所有数据复制处理。另外，对于在复制处理中到达的主机 I/O 请求，进行如下和同步复制的进展对应的 I/O 处理：若对象区域是同步复制结束的区域，则在复制源和复制目的地两方进行双重写入，若对象区域是同步复制未结束的区域，则只在复制源进行写入。关于在存储装置 130 执行的同步复制处理，在后述的磁盘移动处理 222 中进行说明。

从移动源存储装置 130a 接收到上述的同步复制的所有数据复制处理的结束报告后，管理服务器 110 指示主机 100 将对于移动目的地存储装置 130b 的 LU 的路径定义为移动源存储装置 130a 的第一逻辑设备的交替路径（步骤 719）。

其后，管理服务器 110 指示移动源存储装置 130a 进行向第一逻辑设备的、来自一台或其以上的主机 100 的 I/O 停止和同步复制处理的停止（步骤 720），指示一台或其以上的主机 100 进行移动源存储装置 130a 的 LU 交替路径定义解除和第一逻辑设备的初始化（步骤 721、722）。关于在存储装置 130 执行的 I/O 停止的处理，在后述的设备移动处理 222 中进行说明。在本步骤 722 中，作为第一逻辑设备的初始化，既可以消除第一逻辑设备的存储数据，也可以解除与第一逻辑设备对应的物理设备的对应关系，将第一逻辑设备设为“未安装”状态。另外，也可以不消除存储数据，而只解除设定在第一逻辑设备的属性。

管理服务器 110，通过以上的处理，进行与外部设备或内部物理设备对应的逻辑设备的移动处理，向用户或应用程序报告处理完成，结束处理（步骤 723）。

图 8 是表示在存储装置 130 执行的、磁盘移动处理 222 的处理流程的一例的图。存储装置 130 根据来自管理服务器 110 的处理请求内容选择处理（步

骤 801)。

在处理请求是向特定逻辑设备的超高速缓冲存储器直通模式的设定的情况下，存储装置 130 将与对象逻辑设备对应的逻辑设备管理信息 201 的记录 311 设为接通（步骤 802），将关于该逻辑设备在磁盘超高速缓冲存储器 134 中保存的未反映的更新数据都写入到该逻辑设备对应的物理设备或外部设备（步骤 803）。

在来自管理服务器 110 的处理请求是同步复制指示的情况下，根据请求参数决定成为对象的第一逻辑设备和移动目的地存储装置 130b 以及移动目的地第二逻辑设备。另外，在本步骤 805，在对应于第一逻辑设备的逻辑设备管理信息 201 的记录 308 到 310 设定控制信息，以使在和存储装置 130b 的第二逻辑设备之间形成同步复制对（步骤 805）。

然后，使用和第一逻辑设备对应的逻辑设备管理信息 201 的记录 309（数据转移进展指针）的同时，将第一逻辑设备的所有数据按照次序复制到第二逻辑设备（步骤 806）。此外，就算所有数据的复制结束，只要没有被后述的 I/O 停止请求等解除，就继续进行在第一以及第二逻辑设备定义的同步复制。此时，同步复制中接受的对该逻辑设备的、来自主机 100 的数据更新，通过后述的存储装置 I/O 处理 221，和主机请求同步地复制到移动目的地存储装置 130b 的第二逻辑设备。

另一方面，在来自管理服务器 110 的处理请求是对特定逻辑设备的 I/O 停止请求的情况下，由请求参数决定成为对象的逻辑设备和停止 I/O 请求的一台或其以上的主机 100，从与该逻辑设备对应的逻辑设备管理信息 201 的记录 307 中，消除成为 I/O 停止对象的一台或其以上的主机 100 的标识符，停止从该主机 100 对该逻辑设备的 I/O 请求（步骤 807）。

接着，若该逻辑设备在同步复制中，则更新对应的逻辑设备管理信息 201 的记录 308 到 310，以解除同步复制，使其成为通常状态（步骤 808）。

图 9 是表示在存储装置 130 执行的、存储装置 I/O 处理 221 的处理流程的一例的图。经由端口 131 接受来自主机 100 的 I/O 请求的存储装置 130 根据逻辑设备管理信息 201 的记录 307 等信息确认来自该请求的访问源主机 100 的访问权（步骤 901），在来自不具有访问权的主机 100 的请求的情况下，拒

绝该请求（步骤 908）。

在步骤 901 可接受该 I/O 请求的情况下，存储装置 130 检查逻辑设备管理信息 201 的记录 311 的超高速缓冲存储器直通标志，判断是否是超高速缓冲存储器直通模式（步骤 902）。

在是超高速缓冲存储器直通模式、且该请求是写入请求的情况下（步骤 903），存储装置 130 将从主机 100 接受的更新数据存储到磁盘超高速缓冲存储器 134 中（步骤 904）。其后，存储装置 130 将该更新数据写入到与该逻辑设备对应的物理/外部设备（步骤 905），将保存在磁盘超高速缓冲存储器 134 的更新数据变更为反映完毕状态（步骤 906），向主机 100 报告该写入请求处理的结束（步骤 907）。

另外，在成为对象的逻辑设备不是超高速缓冲存储器直通模式（步骤 902）、但在同步复制中且该 I/O 请求是写入请求的情况下，存储装置 130 将从主机 100 接受的更新数据存储到磁盘超高速缓冲存储器 134 中（步骤 912）。其后，存储装置 130 将该更新数据写入到转移目的地存储装置 130b 的第二逻辑设备（步骤 913），向主机 100 报告该写入请求处理的结束（步骤 907）。

另外，在对象逻辑设备是超高速缓冲存储器直通模式且该请求是写入请求以外的情况下，或对象逻辑设备不是超高速缓冲存储器直通模式且不在同步复制中或在同步复制中但该请求是写入请求以外的情况下，存储装置 130 处理被请求的 I/O 请求（步骤 909）。

在步骤 909 中，对于读取请求处理，存储装置 130 检查对象数据是否保存在磁盘超高速缓冲存储器 134 内，如果没有，则从物理设备或外部设备读取，将该数据存储到磁盘超高速缓冲存储器 134 的同时，转发给主机 100。另外，对写入请求，存储装置 130 在将从主机 100 接受的更新数据保存在磁盘超高速缓冲存储器 134 的时刻，向主机 100 报告该写入请求处理结束，以后按适当的定时写入到对应的物理设备或外部设备。

图 10 是表示在主机 100 执行的、主机 I/O 处理 261 的处理流程的一例的图。

主机 100 从在主机 100 上执行的应用程序接受与存储装置 130 的 LU 对应的向特定设备文件的 I/O 请求（步骤 1001）。

主机 100 参照设备路径管理信息 251 检查在该设备文件是否设定了多个路径（多路径）（步骤 1002）。

在没有设定多个路径的情况下，将指定设备文件作为访问对象，在设定的情况下，根据各路径的状态和 I/O 分级的优先顺序，决定成为访问对象的设备文件（步骤 1003）。

接着，主机 100 对作为访问对象决定的设备文件进行变换，计算对应的存储装置 130 以及端口 131 的标识符和 LUN（步骤 1004）。

根据求出来的地址信息，主机 100 发送向对象设备的 I/O 请求，在正常结束时向应用程序报告该 I/O 请求处理结束（步骤 1006、1009）。

在步骤 1006，在异常结束时检查对该设备的交替路径的有无（步骤 1007），若交替路径存在，则通过该路径重发该 I/O 请求，在 I/O 再次尝试成功的情况下，报告该 I/O 请求处理结束（步骤 1009）。

另外，在步骤 1007，在交替路径不存在的情况下，报告该 I/O 请求处理异常结束（步骤 1009）。

## 【实施例 2】

参照图 1 至图 6、图 10、图 12 至图 14 对第二实施例进行说明。由于第一和第二实施例具有很多共同点，因此只对两者的差异进行说明。

在第二实施例中，首先在移动目的地存储装置 130b 定义将移动源存储装置 130a 的第一逻辑设备虚拟化的外部设备，若对应该外部设备定义了第二逻辑设备，则将一台或其以上的主机 100 的访问对象切换至第二逻辑设备。其后，将停留在移动源存储装置 130a 的磁盘超高速缓冲存储器 134 内的未反映的更新数据写入到外部存储装置 150 的第三逻辑设备后，在移动目的地存储装置 130b 定义将第三逻辑设备虚拟化的别的外部设备，并切换对应关系，以使第二逻辑设备与后者的外部设备对应起来。

在第一和第二实施例中，图 1 所示的硬件结构以及图 2 所示的软件结构基本相同。但是，由于设备切换顺序的不同，在设备移动指示处理 241、设备移动处理 222、存储装置 I/O 处理 221 的三个处理中有差异。

图 12 是表示在第二实施例中的管理服务器 110 执行的、设备移动指示处理 241 的处理流程的一例的图。和第一实施例一样，接受来自用户或管理者

或应用程序的设备移动请求的管理服务器 110，根据移动请求的参数信息，首先在移动目的地存储装置 130b 定义将移动源存储装置 130a 的第一逻辑设备虚拟化的第二外部设备。同时，在本步骤，管理服务器 110 定义与第二外部设备对应的第二逻辑设备（步骤 1201）。

然后，管理服务器 110，将在移动源存储装置 130a 中设定在第一逻辑设备的属性，设定在移动目的地存储装置 130b 的第二逻辑设备（步骤 1202），在移动目的地存储装置 130b 进行向第二逻辑设备的 LU 路径定义（步骤 1203），使一台或其以上的主机 100 识别该 LU（步骤 1204）。

接着，管理服务器 110 指示移动目的地存储装置 130b 设定第二逻辑设备的超高速缓冲存储器直通模式（步骤 1205）。

在该状态下，管理服务器 110 指示一台或其以上的主机 100，将移动目的地存储装置 130b 的 LU 定义为对移动源存储装置 130a 的对象 LU 的路径的交替路径（步骤 1206），进而，对于移动源存储装置 130a，指示对第一逻辑设备的来自一台或其以上的主机 100 的 I/O 停止（步骤 1207）。

这样，由于拒绝来自相应的一台或其以上的主机 100 的、向移动源存储装置 130a 的 LU 的 I/O 请求，因此向设定了交替路径的移动目的地存储装置 130b 的 LU 发送其后的 I/O 请求。

进而，管理服务器 110，指示一台或其以上的主机 100 从交替路径中解除对移动源存储装置 130a 的 LU 的路径，指示移动目的地存储装置 130b 解除第二逻辑设备的超高速缓冲存储器直通模式（步骤 1208、1209）。

这样，能够将一台或其以上的主机 100 的访问对象切换至移动目的地存储装置 130b 的 LU。在第二实施例中，从该状态进一步实施将移动源存储装置 130a 的第一逻辑设备移动至第二逻辑设备的处理。

首先，管理服务器 110 判定成为移动对象的第一逻辑设备对应于外部设备或移动源存储装置 130a 内部的物理设备中的哪一个（步骤 1210）。

在对应于外部设备的情况下，管理服务器 110 指示设定移动源存储装置 130a 的第一逻辑设备的超高速缓冲存储器直通模式，并保存在磁盘超高速缓冲存储器 134 中，将向第一逻辑设备（即，外部存储装置 150 的第三逻辑设备）的未反映的更新数据都写入到第三逻辑设备（步骤 1211）。

若未反映更新数据的写入全部结束了，管理服务器 110，对外部存储装置 150 内的第三逻辑设备变更访问权，以使可进行来自移动目的地存储装置 130b 的访问（步骤 1212），在移动目的地存储装置 130b 定义将第三逻辑设备虚拟化的第三外部设备（步骤 1213）后，变更移动目的地存储装置 130b 中的第二逻辑设备和第二外部设备之间的对应，使第二逻辑设备和第三外部设备对应起来（步骤 1214）。

然后，变更访问权，以使禁止从移动源存储装置 130a 访问外部存储装置 150 的第三逻辑设备（步骤 1216）。

这些处理结束后，管理服务器 110 向请求源报告该设备移动请求处理结束（步骤 1217）。

另一方面，在第一逻辑设备对应于物理设备的情况下，从移动目的地存储装置 130b 内的物理设备中，选定不对应于逻辑设备的未使用的第二物理设备，使第二逻辑设备和第二物理设备对应（步骤 1218）。

然后，指示移动目的地存储装置 130b 将第一逻辑设备的全部数据转移到第二物理设备（步骤 1219）。

若通过来自存储装置 130b 的结束报告检测到全部数据的转移已结束，管理服务器 110 向请求源报告该处理结束（步骤 1217）。

图 13 是表示在第二实施例中、在存储装置 130 执行的设备移动处理 222 的处理流程的一例的图。存储装置 130 检查来自管理服务器 110 的请求处理内容，在超高速缓冲存储器直通模式设定处理的情况下，和第一实施例一样，进行向该逻辑设备的超高速缓冲存储器直通模式设定、以及保存在磁盘超高速缓冲存储器 134 内的对该逻辑设备的未反映更新数据的向物理/外部设备的写入（步骤 1302、1303）。

在超高速缓冲存储器直通解除指示处理的情况下，存储装置 130 将被指定的逻辑设备的逻辑设备管理信息 201 的记录 311（超高速缓冲存储器直通标志）设定为断开（步骤 1305）。

在数据转移处理的情况下，存储装置 130 对由参数信息指定的对象逻辑设备，在对应于该逻辑设备的转移源物理/外部设备和另指定的转移目的地物理/外部设备之间实施数据转移处理，变更逻辑设备和物理/外部设备的对应关

系，以使在全部的数据转移结束的时刻使该逻辑设备与转移目的地物理/外部设备对应起来（步骤 1306、1307）。具体地说，存储装置 130 在与该逻辑设备对应的逻辑设备管理信息 201 的记录 303 中设定转移目的地物理/外部设备号码，将记录 310 的数据转移中标志设为断开。另外，存储装置 130 更新与成为转移目的地以及转移源的物理/外部设备对应的物理/外部设备管理信息 203/204 的记录 503/603 以及记录 504/604。

最后，在请求处理是 I/O 停止处理的情况下，和第一实施例一样，存储装置 130 更新控制信息，以使抑制向对象逻辑设备的、来自被指定的一台或其以上的主机 100 的 I/O 请求（步骤 1308）。

图 14 是表示在第二实施例中、在存储装置 130 执行的存储装置 I/O 处理 221 的处理流程的一例的图。此外，由于步骤 1401 到步骤 1408 的处理，和图 9 所示的第一实施例中的存储装置 I/O 处理 221 的步骤 901 到步骤 908 相同，因此省略其说明。

在第二实施例中，在成为对象的逻辑设备不是超高速缓冲存储器直通模式的情况下，存储装置 130，在该逻辑设备在数据转移中且该 I/O 请求的访问地址为转移完毕区域时，对于转移目的地、在不是转移完毕区域时对转移源、即对当前时刻与该逻辑设备对应的物理/外部设备进行该 I/O 请求处理（步骤 1409 到 14012）。

根据本发明，就可以做到：在利用了外部存储装置连接技术的多个虚拟化控制存储装置中、从第 1 虚拟化控制存储装置向另外的第 2 虚拟化控制存储装置，一边受理针对已被虚拟化的第 3 存储装置的第 3 设备的主机 I/O 请求，一边切换虚拟设备控制。

即，根据由于主机的 I/O 负载变动等而产生的在虚拟化控制存储装置之间负载不均衡的发生，通过在虚拟化控制存储装置之间切换虚拟设备控制的担当，可以实现虚拟化控制存储装置之间的输入输出处理量的均匀化、最佳化。这样，能够具备系统运用中的负载重新调整的机构，能够更简单地进行装载了大量设备的大规模虚拟存储装置系统的性能设计。

在如上所述的第一以及第二实施例中，存储装置 130 之间的设备移动指示，是由用户或管理者或应用程序指定成为移动对象的逻辑设备以及移动目

的地存储装置 130。这里，参照图 11，作为移动对象设备以及移动目的地存储装置 130 的选定处理的一例，对在管理服务器 110 执行的存储装置负载监视处理 242 进行说明。

管理服务器 110 经由 IP 网络 175 周期性地取得存储装置 130 的构成信息，检查在各存储装置 130 是否发生了构成异常（步骤 1101、1102）。

在检测到由于磁盘超高速缓冲存储器 134 或控制处理器 132 的故障而造成特定存储装置 130 的 I/O 处理能力下降的情况下，管理服务器 110 将该存储装置 130 决定为移动源存储装置 130a（步骤 1108）。

在没有检测到相应的存储装置 130 的情况下，管理服务器 110 周期性地从各存储装置 130 取得各逻辑设备的 I/O 频度、各控制处理器 132 或磁盘超高速缓冲存储器 134 的利用率等运行信息（步骤 1103）。

在各存储装置 130 之间的 I/O 负载的差为规定水平或其以上的情况下，管理服务器 110 将高负载存储装置决定为移动源存储装置 130a（步骤 1104、1105）。

对通过步骤 1105 或步骤 1108 决定的移动源存储装置 130a，管理服务器 110 选定成为移动源的第一逻辑设备，进一步选定移动目的地存储装置 130b 以及成为移动目的地的第二逻辑设备，向设备移动指示处理 241 请求该设备移动（步骤 1106、1107）。在本步骤 1106，能够适用这样的选定逻辑：即作为成为移动源的第一逻辑设备，选定正常负载高的逻辑设备等、以将存储装置内的设备间负载均衡化为目的的数据转移中使用的传统上众所周知的选定逻辑。

在本发明中，根据逻辑设备是对应于外部设备还是对应于物理设备，移动处理不同。在对应于外部设备的情况下，只要将停留在超高速缓冲存储器内的未反映更新数据写入到设备就可以，不需要像对应于物理设备的情况那样复制保存在该物理设备的全部数据。因此，在步骤 1106 的成为移动源的第一逻辑设备的选定中，希望优先选择对应于外部设备的逻辑设备。另外，在移动源的第一逻辑设备对应于外部设备的情况下，希望从对包含该外部设备的外部存储装置 150 可连接外部设备的存储装置 130 中，选定移动目的地存储装置 130b。

如上所述，通过存储装置负载监视处理 242，在由于机器故障等不能使

用虚拟化控制存储装置内的超高速缓冲存储器或控制处理器等的资源的情况下，或虚拟化控制存储装置的 I/O 处理能力下降的情况下，能够检测这些状态。另外，检测的结果，将该虚拟化控制存储装置的虚拟设备控制向其他虚拟化控制存储装置调换，能够最大限度地运用存储装置系统整体的 I/O 处理能力。

此外，本发明并不限于上述的实施例，可以有多种变形例。

在本发明的第一以及第二实施例中，管理服务器 110，进行一台或其以上的主机 100、移动源存储装置 130a、以及移动目的地存储装置 130b 之间的转移处理的指示，在管理服务器 110 的控制下，执行设备移动处理。移动源存储装置 130a 和移动目的地存储装置 130b 之间的属性交接或超高速缓冲存储器直通模式的设定、解除的定时都由管理服务器 110 来控制。但是，将这些处理的一部分不是由管理服务器 110 而是由移动目的地存储装置 130b 来进行也可以。例如，移动目的地存储装置 130b 从管理服务器 110 接受设备移动指示，其后，可以通过两个存储装置之间的通信，来实现来自移动源存储装置 130a 的设备属性的取得、或向移动源存储装置 130a 的超高速缓冲存储器直通模式的设定指示等等。

另外，在本发明的实施例中，使用了外部存储装置连接技术的基于存储装置 130 的虚拟存储装置系统为对象，但是在基于具有存储装置虚拟化技术的虚拟开关的虚拟存储装置系统中也可以使用大致相同的设备移动方法。

图 15 表示成为本发明的变形例的硬件结构图。图中的光纤通道开关 120 和控制装置 180a 以及 180b（统称为控制装置 180）协同工作，提供基于存储装置虚拟化技术的虚拟设备。此时，考虑从将外部存储装置 150 虚拟化的控制装置 180a 向不同的控制装置 180b 的设备移动。在本发明的变形例中，和在本发明的第一或第二实施例中说明的基于存储装置 130 的虚拟存储装置系统不相同，各控制装置 180 不装载磁盘超高速缓冲存储器 134，因此不需要进行停留在磁盘超高速缓冲存储器 134 内的移动源逻辑设备的更新数据的写入。作为一例，图 16 表示在管理服务器 110 执行的设备移动指示处理 241 的处理流程图。在本发明的变形例中，也可以不考虑停留在磁盘超高速缓冲存储器内的更新数据，因此仅通过在设备间的属性交接，就能够实现在控制装置 180 之间的设备移动。

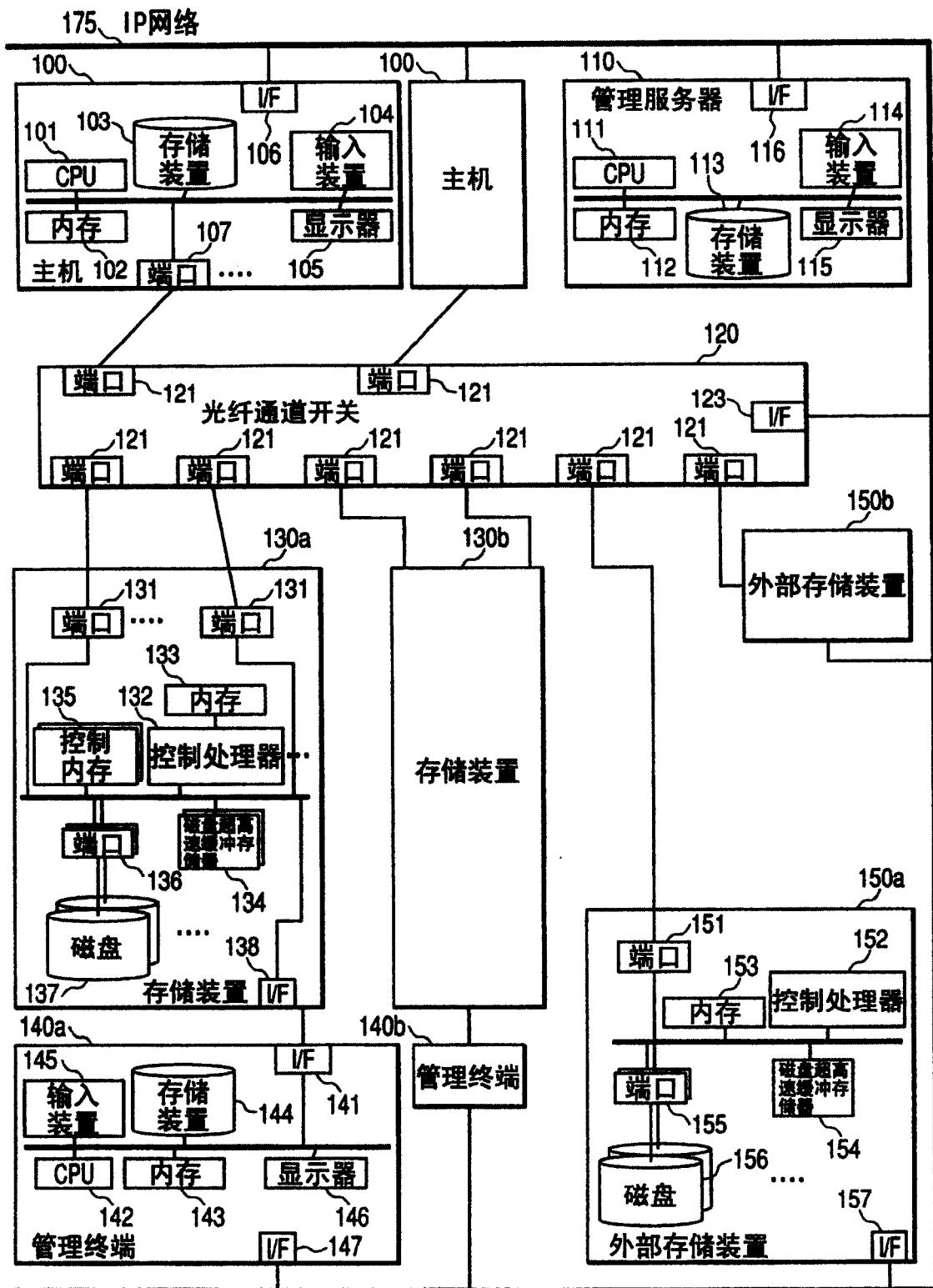


图1

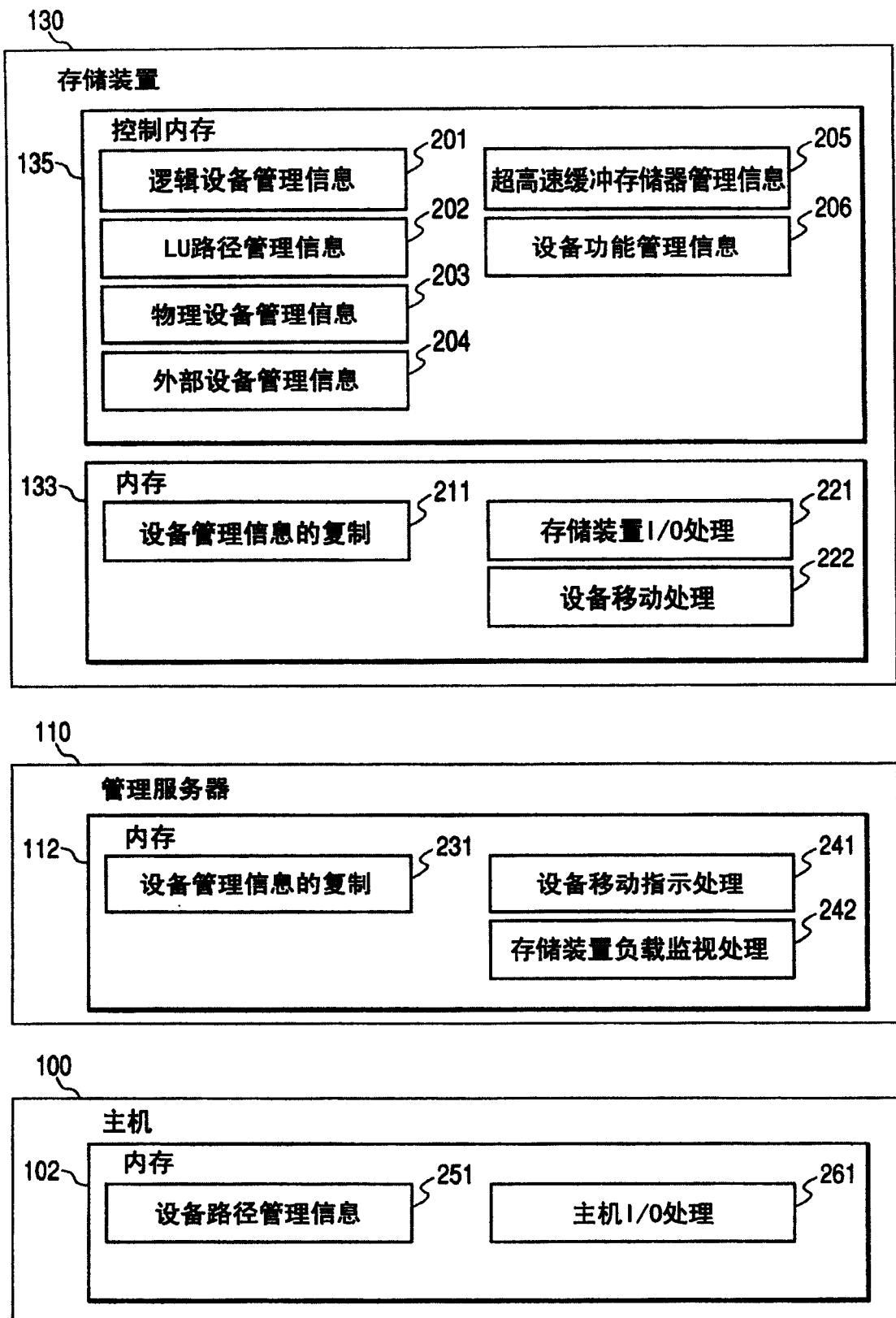


图2

图3

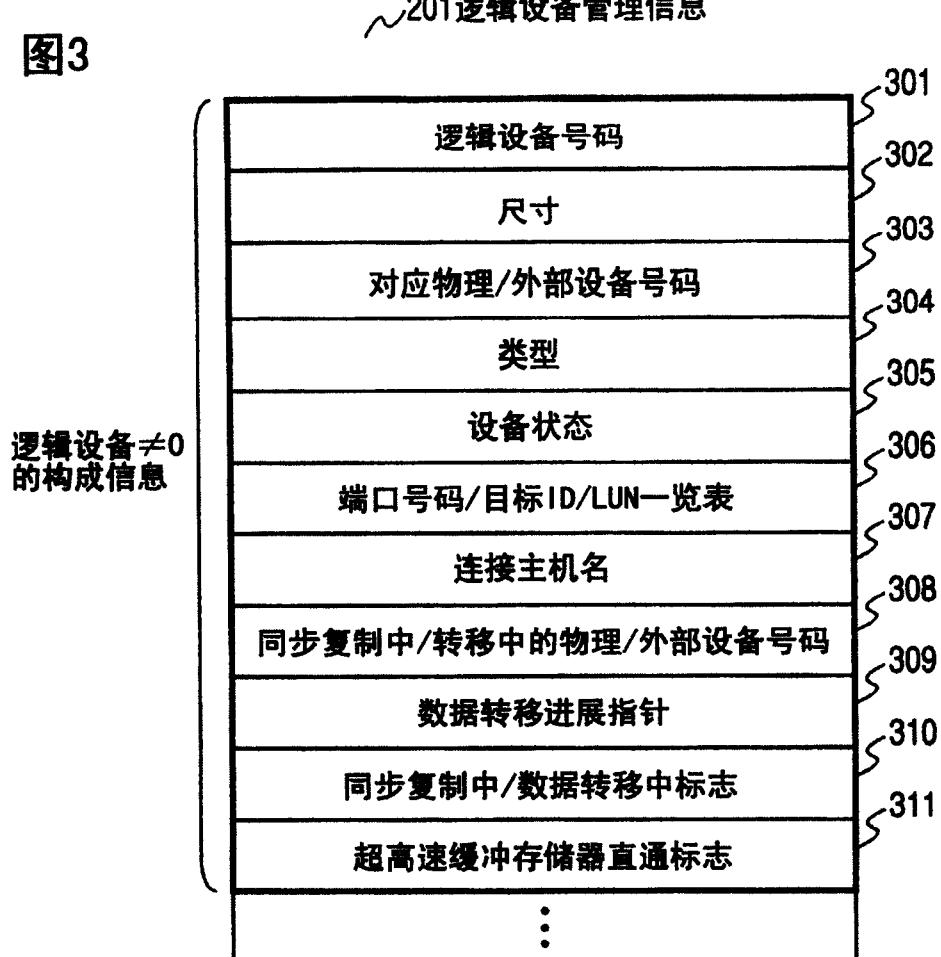
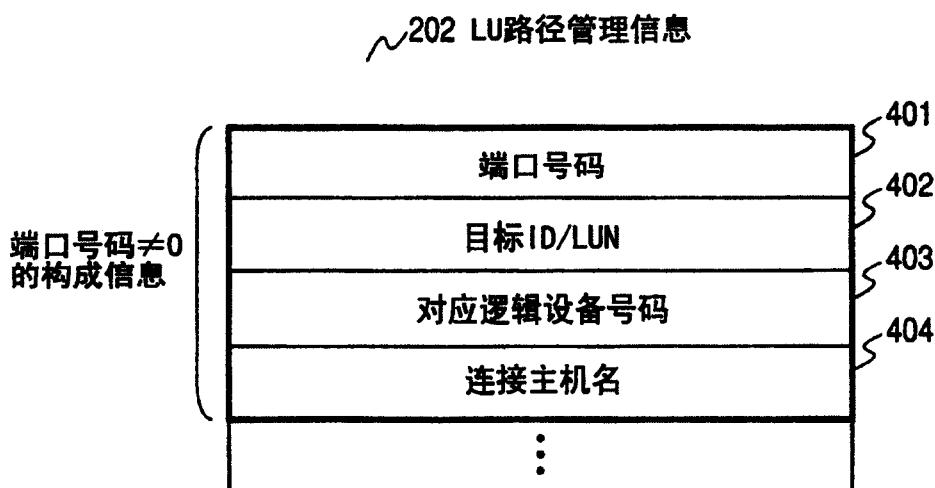
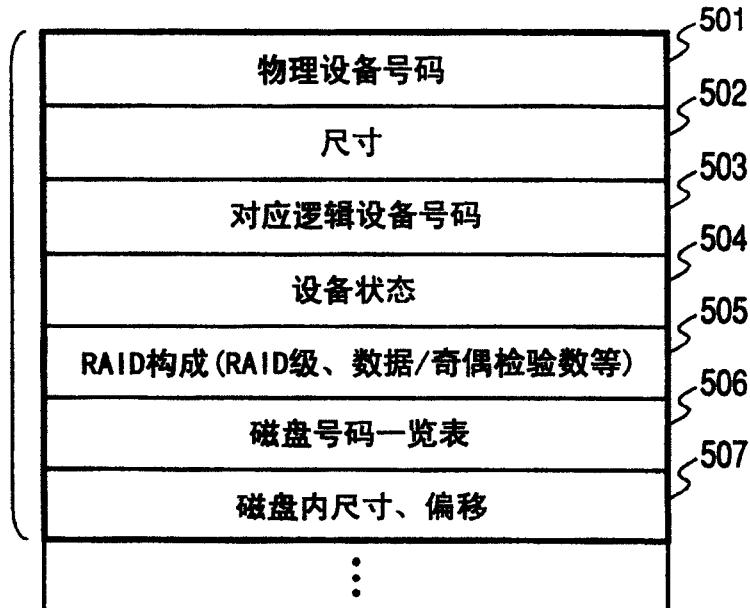


图4



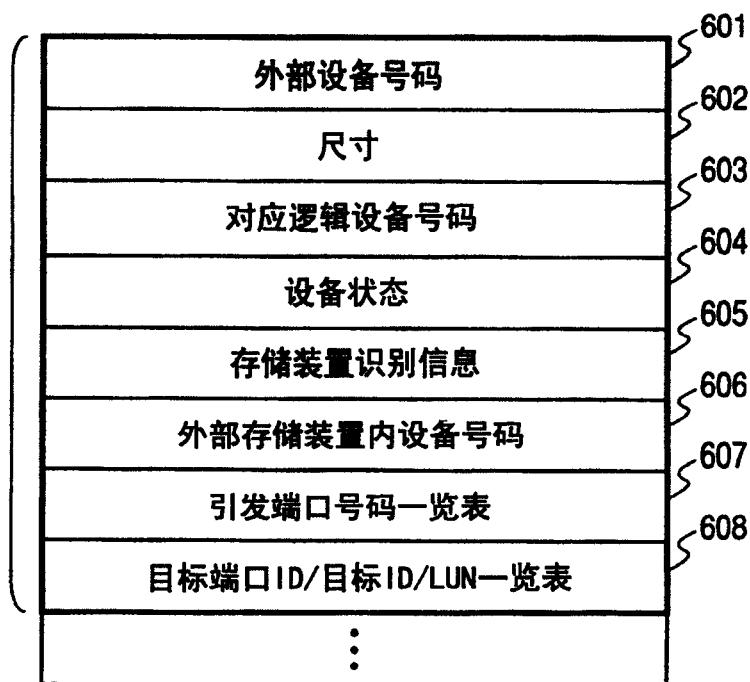
## ~203 物理设备管理信息

图5

物理设备≠0  
的构成信息

## ~204 外部设备管理信息

图6

外部设备≠0  
的构成信息



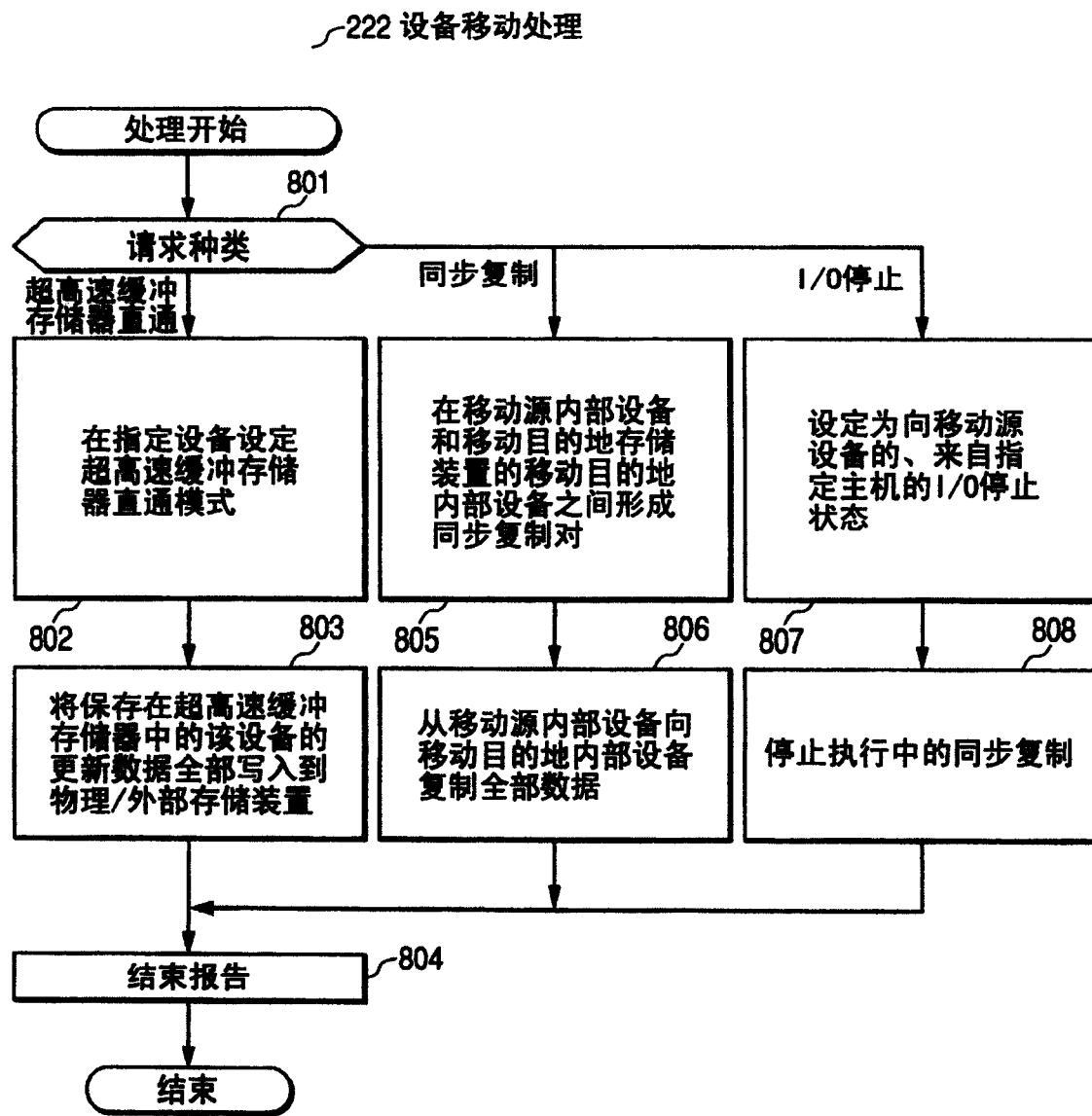


图8

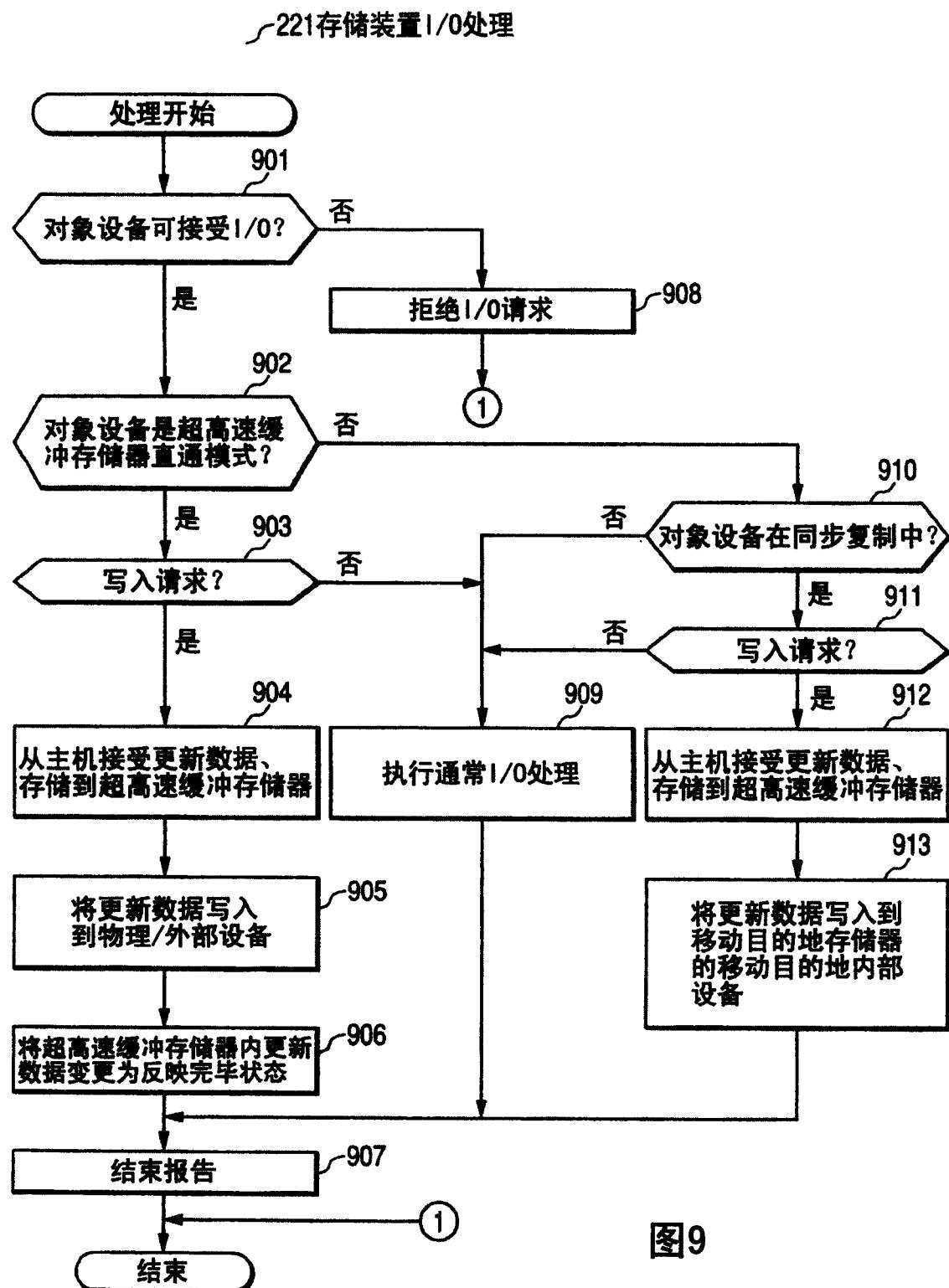


图9

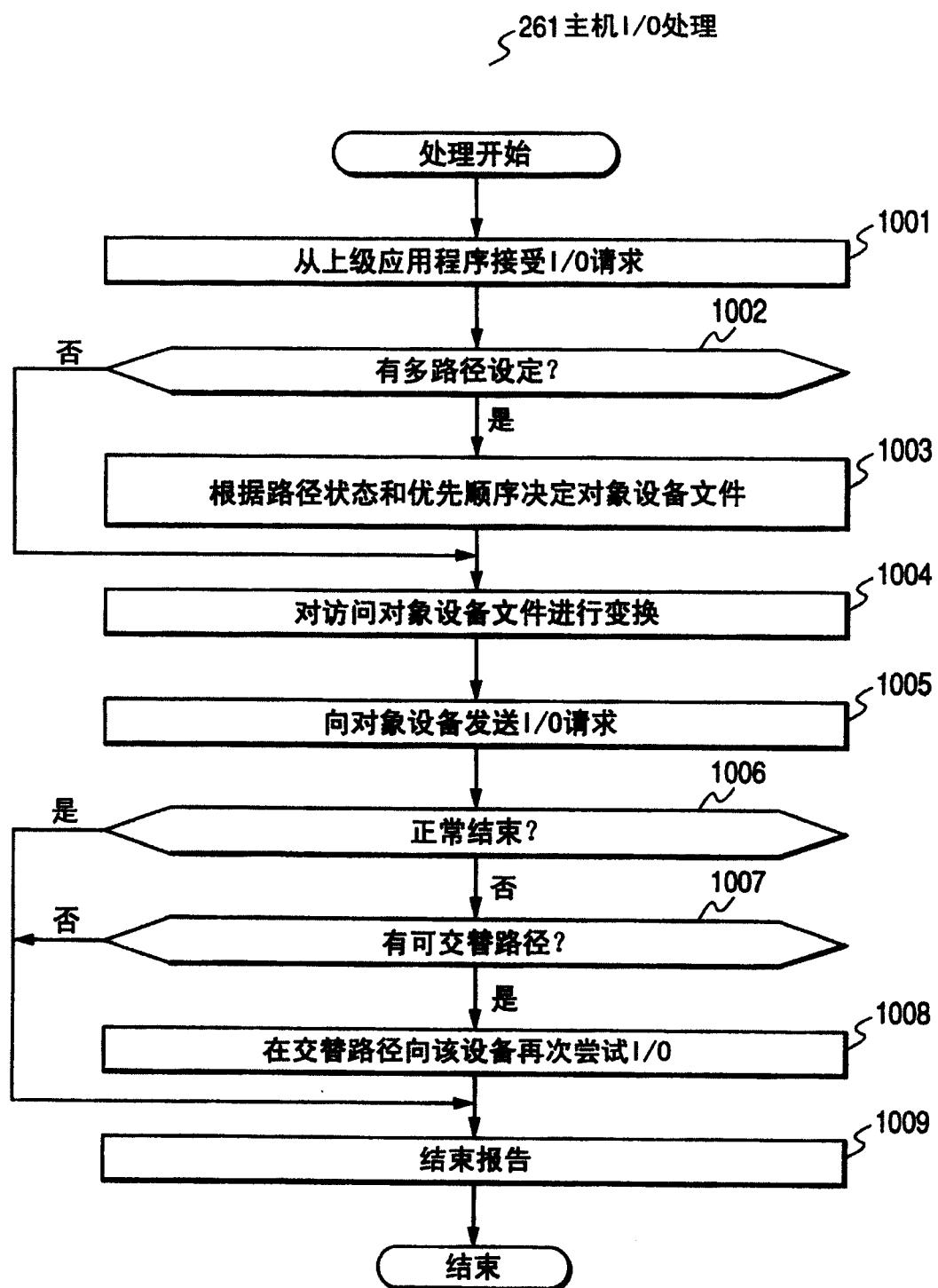


图10

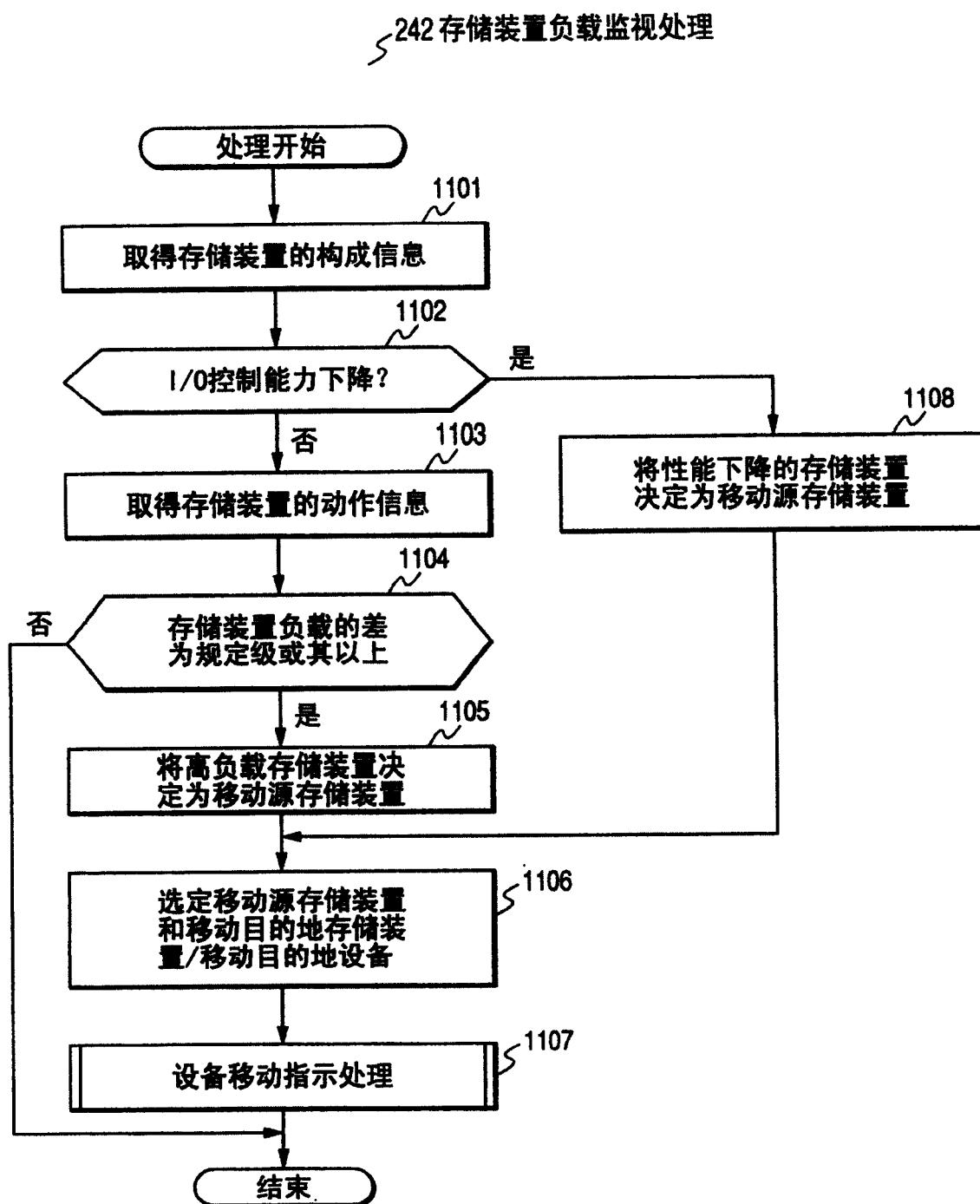


图11

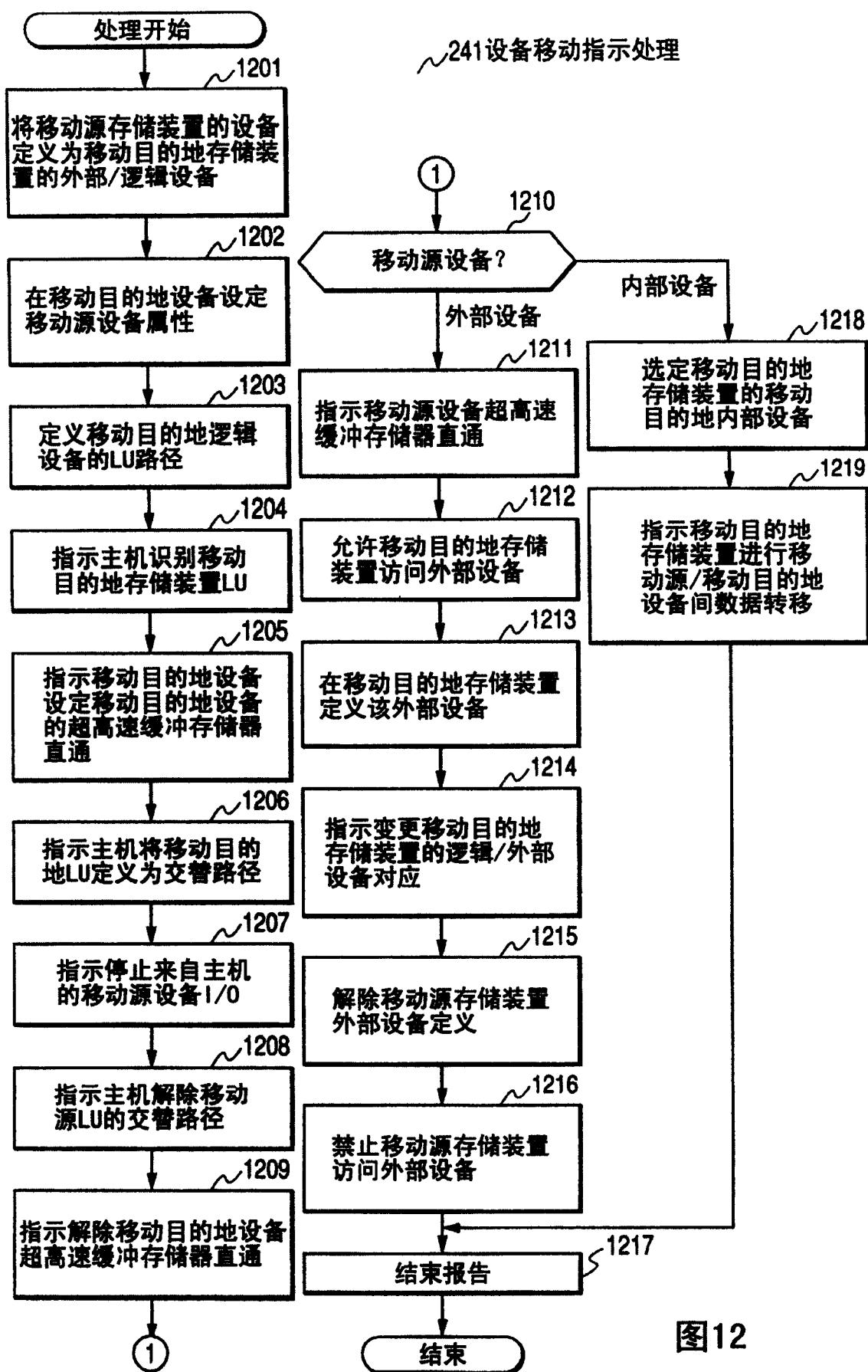


图12

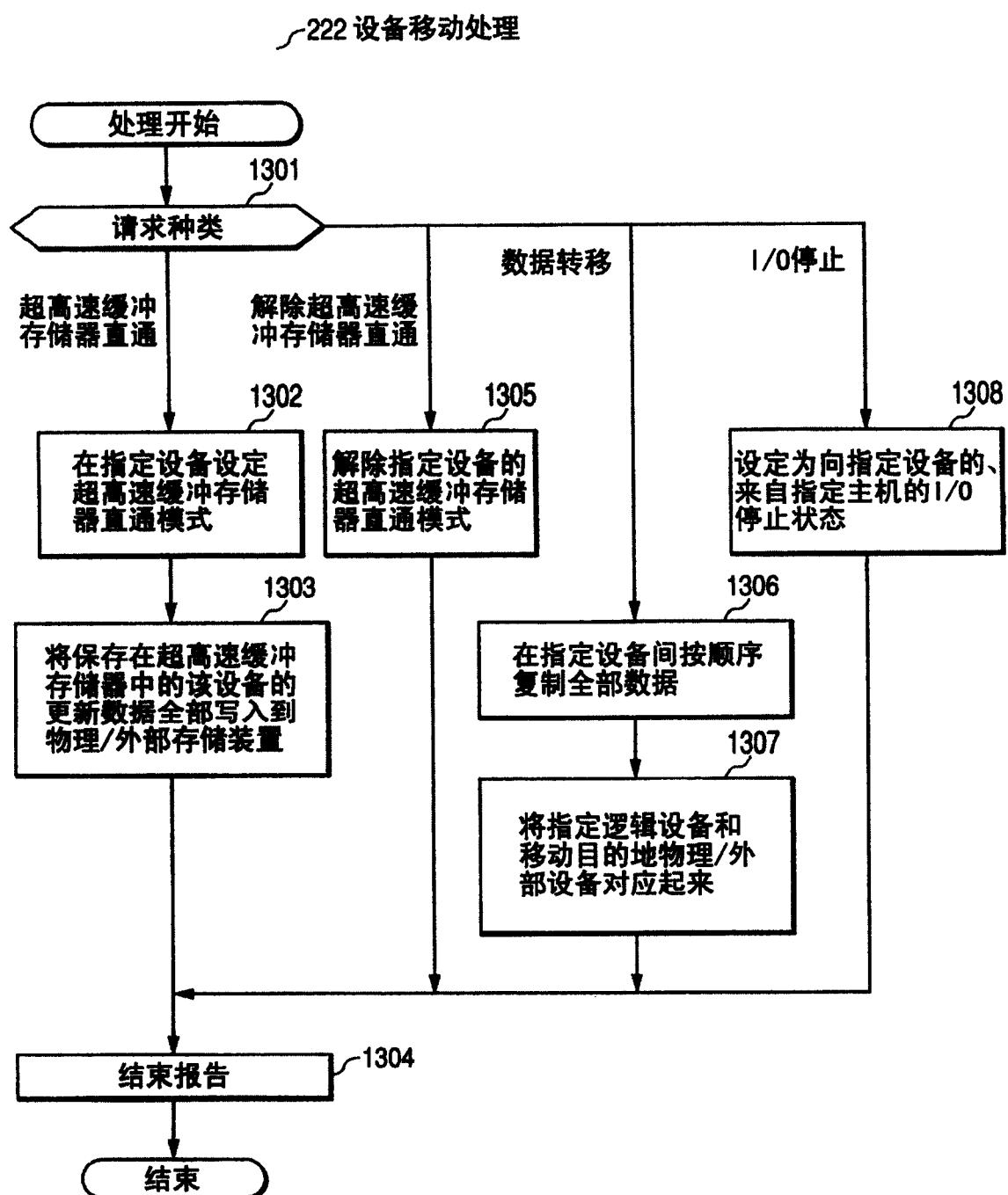


图13

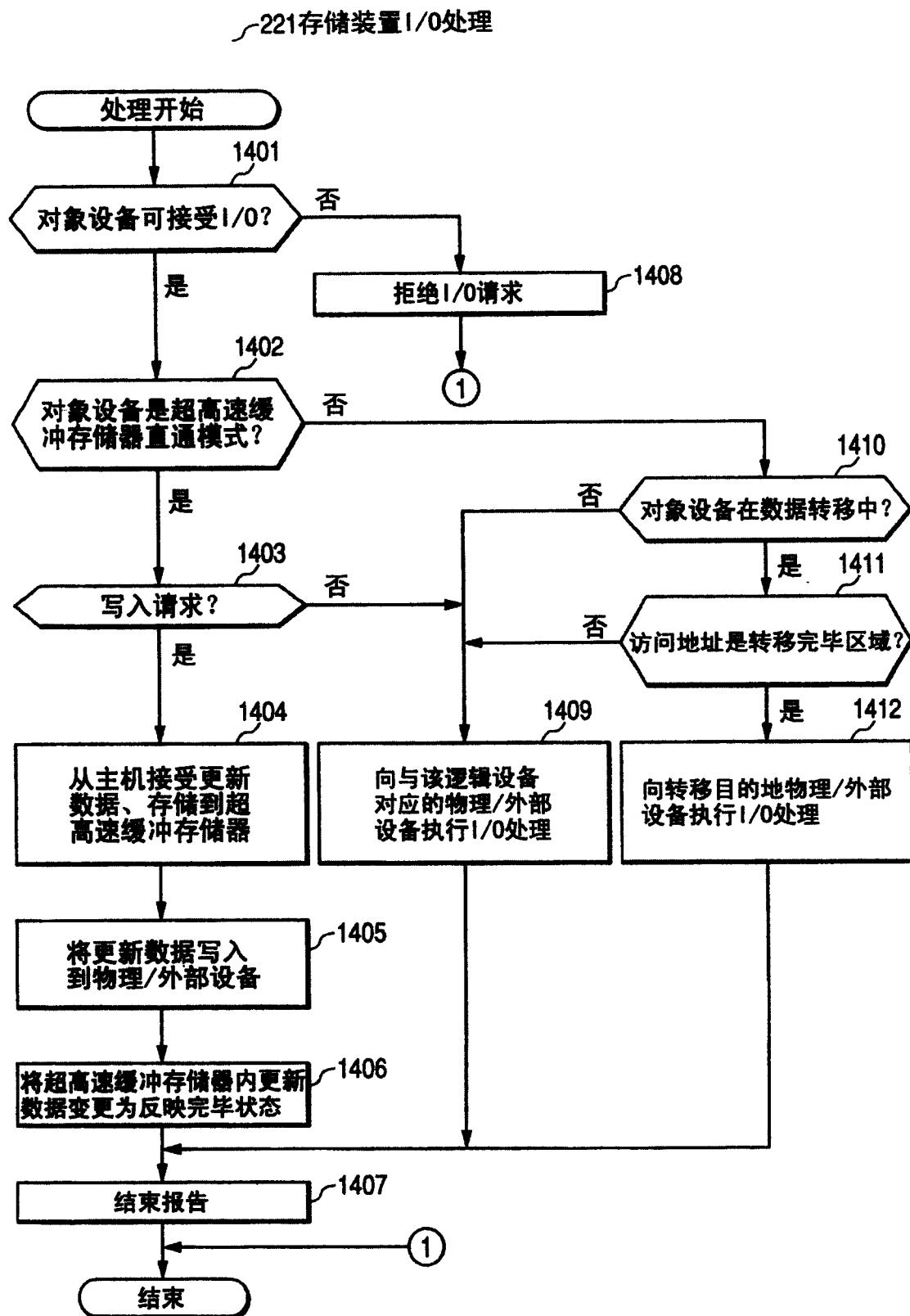


图14

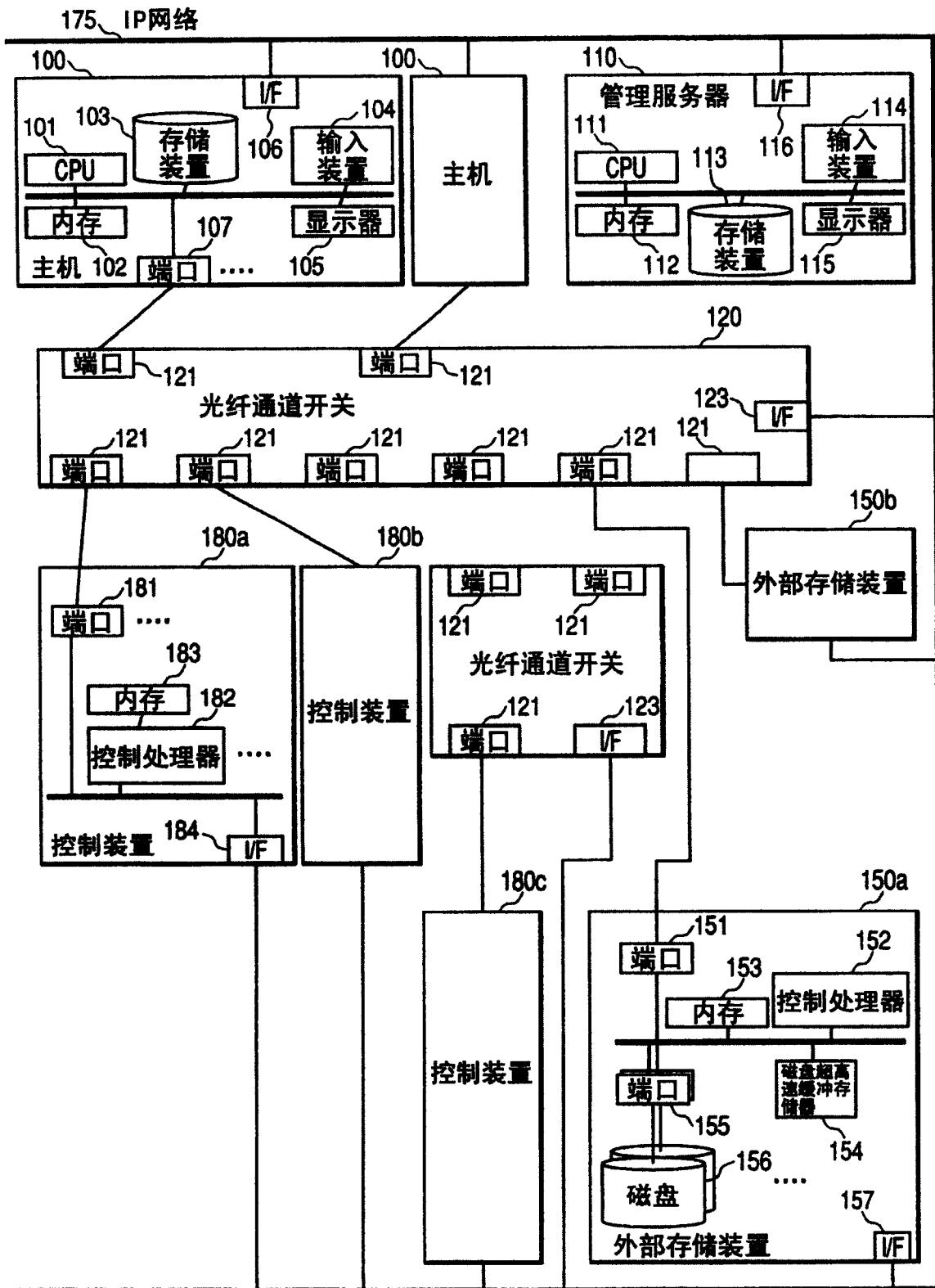


图15

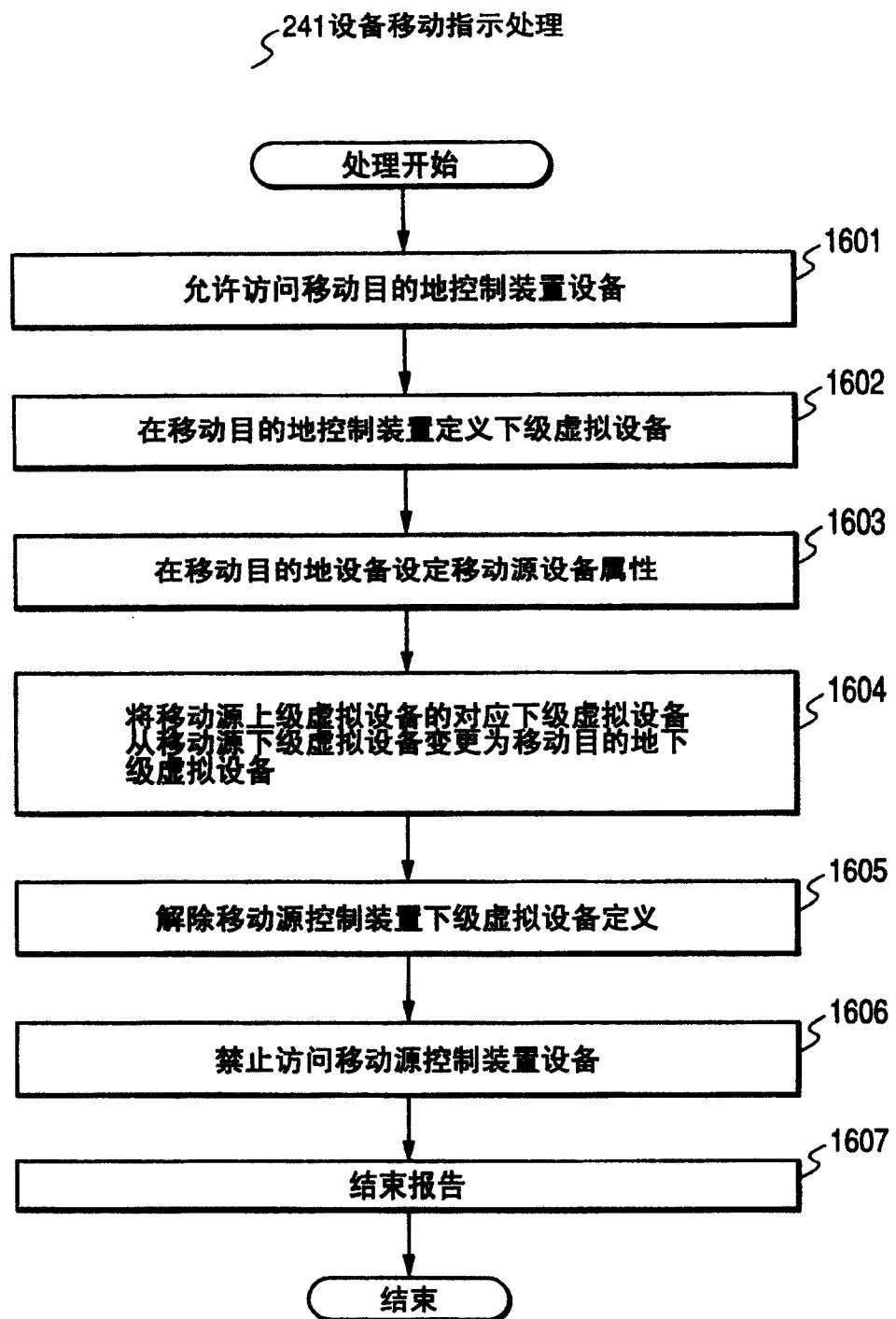


图16