[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 3/06 (2006.01)

[12] 发明专利说明书

专利号 ZL 200410029465.6

[45] 授权公告日 2007 年 8 月 29 日

[11] 授权公告号 CN 100334535C

[22] 申请日 2004.3.19
[21] 申请号 200410029465.6
[30] 优先权
[32] 2003.9.16 [33] JP [31] 2003-323120
[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 桧垣诚一 岛田朗伸 冈见吉规 中野俊夫
[56] 参考文献
CN1387125A 2002.12.25
US5720028A 1998.2.17
US6289398B1 2001.9.11
US6567865B1 2003.5.20
审查员 郑 红

[74] 专利代理机构 北京银龙知识产权代理有限公司
代理人 韩惠琴

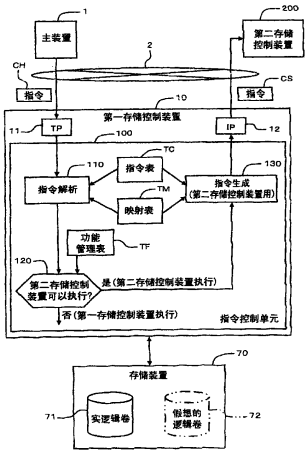
权利要求书 5 页 说明书 33 页 附图 22 页

[54] 发明名称

存储系统及存储控制装置

[57] 摘要

本发明提供一种存储系统及存储控制装置，该系统及装置在逻辑上汇集并假想地提供分散在多个存储控制装置上的存储区域，通过特定的处理使数据处理分散，以便不集中在一个存储控制装置上。第一存储控制装置(10)具备把第二存储控制装置(200)所具有的逻辑卷当作自己的卷的假想的LU72。 在从主装置(1)请求的数据处理是例如负荷大的特定的处理(直接复制，逻辑卷复制)的场合，根据功能管理表，判断第二存储控制装置能否执行请求的处理。 在能够执行的场合，向第二存储控制装置发送指令，使其代行处理。



1．存储系统，所述系统包括第一存储控制装置和第二存储控制装置，第一存储控制装置和第二存储控制装置连接、且第一存储控制装置和第二存储控制装置能够进行通信，是进行响应来自上位装置的请求的数据处理的存储系统，其特征在于，

所述第一存储控制装置具有第一控制装置，该第一控制装置判断第二存储控制装置能否执行从所述上位装置接受的第一请求所涉及的规定的数据处理，在判断为所述第二存储控制装置能够执行的场合，生成与所述第一请求对应的第二请求，并发送给所述第二存储控制装置；

所述第二存储控制装置具有第二控制装置，该第二控制装置根据从所述第一存储控制装置接收的所述第二请求，进行所述规定的数据处理；

其中所述第一存储控制装置是把所述第二存储控制装置管理的第二存储区域作为自己管理的第一存储区域假想地提供给上述上位装置的装置；

所述第一请求是请求涉及所述第一存储区域的数据处理的请求。

2．如权利要求 1 所述的存储系统，其特征在于，

所述第一存储控制装置是保持表示所述第一存储区域和所述第二存储区域的对应关系的存储区域对应信息，根据该存储区域对应信息向所述上位装置假想地提供第一存储区域的装置；

所述第一控制装置能够根据所述存储区域对应信息，执行通过所述第一请求请求的以第一存储区域为对象的数据处理。

3．如权利要求 1 所述的存储系统，其特征在于，所述第二请求具有与所述第一请求同样的数据结构。

4．如权利要求 1 所述的存储系统，其特征在于，所述第一控制装置在将所述第二请求发送给所述第二存储控制装置前，确认所述第二存储控制装置能否执行所述第二请求涉及的规定的数据处理。

5．如权利要求 1 所述的存储系统，其特征在于，所述第一存储控制装置保持表示在所述第二存储控制装置能够执行的数据处理功能的功能

管理信息；

所述第一控制装置根据所述功能管理信息，判断能否在所述第二存储控制装置执行所述第二请求涉及的规定的数据处理。

6．如权利要求 5 所述的存储系统，其特征在于，所述功能管理信息是在定义存储系统的构成时，手动或者自动地生成的信息。

7．如权利要求 1 所述的存储系统，其特征在于，

具备备份装置，该备份装置与所述第一存储控制装置以及所述第二存储控制装置两者连接并可以进行通信；

所述第一控制装置是如下装置，即，在所述第一请求所涉及的数据处理是要把存储在所述第一存储区域中的信息转发到所述备份装置并进行存储的备份处理的场合，判断所述第二存储控制装置能否执行所述备份处理，在判断为所述第二存储控制装置能够执行所述备份处理的场合，通过将包含在所述第一请求中的所述第一存储区域中的地址变换为所述第二存储区域中的地址，生成所述第二请求，并发送给所述第二存储控制装置；

所述第二控制装置，根据所述第二请求，把存储在所述第二存储区域中的信息转发到所述备份装置并存储。

8．如权利要求 1 所述的存储系统，其特征在于，

所述第一存储控制装置是把与第一存储区域成对的第一辅助存储区域进一步假想地提供的装置；

所述第二存储控制装置进一步具有与第二存储区域成对的第二辅助存储区域；

所述第一控制装置是如下装置，即，在所述第一请求所涉及的规定的数据处理是把存储在所述第一存储区域中的信息复制到所述第一辅助存储区域的内部复制处理的场合，判断所述第二存储控制装置能否执行所述内部复制处理，在判断为所述第二存储控制装置能够执行所述内部复制处理的场合，通过将包含在所述第一请求中的所述第一存储区域中的地址变换为所述第二存储区域中的地址，生成所述第二请求，发送给所述第二存储控制装置；

所述第二控制装置，根据所述第二请求把存储在所述第二存储区域中的信息复制到所述第二辅助存储区域中。

9．如权利要求 1 所述的存储系统，其特征在于，

具有和主场所连接并可以进行通信的副场所，该副场所和包括第一存储控制装置以及第二存储控制装置设置的主场所成对；

所述副场所具备另外的第一存储控制装置以及另外的第二存储控制装置，所述另外的第一存储控制装置是把所述的另外的第二存储控制装置管理的另外的第二存储区域作为自己管理的另外的第一存储区域假想地提供的装置；

所述主场所的第一控制装置，在所述第一请求所涉及的规定的数据处理是要把存储在第一存储区域中的信息复制到所述副场所的另外的第一存储区域的外部复制处理的场合，判断所述第二存储控制装置以及所述另外的第二存储控制装置的双方能否执行所述外部复制处理，在判断为所述第二存储控制装置和所述另外的第二存储控制装置分别能够执行所述外部复制处理的场合，生成与所述第一请求对应的第二请求，发送给所述第二存储控制装置；

所述第二控制装置，根据所述第二请求，通过使存储在所述第二存储区域中的信息复制到所述另外的第二存储区域中，执行所述外部复制处理。

10．如权利要求 9 所述的存储系统，其特征在于，

所述另外的第一存储控制装置保持表示在所述另外的第二存储控制装置能够执行的数据处理功能的另外的功能管理信息；

所述第一控制装置是在把所述第二请求发送给所述第二存储控制装置之前，通过向所述另外的第一存储控制装置问询，判断所述另外的第二存储控制装置能否执行所述外部复制处理。

11．如权利要求 9 所述的存储系统，其特征在于，

所述第一存储控制装置保持表示所述第一存储区域和所述第二存储区域的对应关系的存储区域对应信息；

所述另外的第一存储控制装置保持表示所述另外的第一存储区域和

所述另外的第二存储区域的对应关系的另外的存储区域对应信息；

所述第一控制装置，在将所述第二请求发送给所述第二存储控制装置的场合，发送所述第一存储控制装置保持的存储区域对应信息和所述的另外的第一存储控制装置保持的另外的存储区域对应信息。

12．如权利要求9所述的存储系统，其特征在于，

所述第一存储控制装置具备更新位置信息保持装置，该更新位置信息保持装置保持涉及在所述外部复制处理中被所述上位装置更新过的所述第一存储区域的信息的信息；

所述第一控制装置，在所述外部复制处理结束的场合，为了使在所述第一存储区域中更新过的信息存储在所述另外的第一存储区域中，根据所述更新位置信息保持装置，生成所述第二请求，从所述第二存储控制装置读出更新过的信息，发送所述读出的信息。

13．存储系统的控制方法，该方法是用于控制包括第一存储控制装置和第二存储控制装置，其中第一存储控制装置和第二存储控制装置连接且能够进行通信，进行响应来自上位装置的请求的数据处理的存储系统的控制方法，其特征在于，

所述第一存储控制装置执行：从所述上位装置接收第一请求的步骤；判断所述第二存储控制装置能否执行所述接收的第一请求所涉及的规定的数据处理的步骤；在判断为所述第二存储控制装置能够执行的场合，生成与所述第一请求对应的第二请求的步骤；把所述生成的第二请求发送给所述第二存储控制装置的步骤；其中在所述的第一存储控制装置执行中，所述第一存储控制装置是把所述第二存储控制装置管理的第二存储区域作为自己管理的第一存储区域假想地提供给上述上位装置的装置；

所述第一请求是请求涉及所述第一存储区域的数据处理的请求；

所述第二存储控制装置执行：从所述第一存储控制装置接收所述第二请求的步骤；和根据所述接收的第二请求进行所述规定的数据处理的步骤。

14．存储控制装置，所述存储控制装置是连接另一存储控制装置以

及上位装置并能够与其进行通信，进行响应来自所述上位装置的请求的数据处理的存储控制装置，其特征在于，包括：

接收装置，该接收装置接收来自所述上位装置的第一请求；判断装置，该判断装置判断所述另一存储控制装置能否执行所述接收的第一请求所涉及的规定的数据处理；请求装置，该请求装置在判断为所述另一存储控制装置能够执行的场合生成与所述第一请求对应的第二请求；和发送装置，该发送装置将所述生成的第二请求发送给所述另一存储控制装置；

其中所述连接另一存储控制装置的存储控制装置是把所述另一存储控制装置管理的第二存储区域作为自己管理的第一存储区域假想地提供给上述上位装置的装置；

所述第一请求是请求涉及所述第一存储区域的数据处理的请求。

存储系统及存储控制装置

技术领域

本发明涉及存储系统以及存储控制装置。

背景技术

例如，在处理如数据中心等大规模的数据的数据库系统中，使用和主机分开构成的存储系统管理数据。这种存储系统由例如磁盘阵列装置等构成。磁盘阵列装置是把多个磁盘存储装置配置成阵列状而构成的装置，例如基于 RAID（Redundant Array of Independent Inexpensive Disks）而构建。在磁盘装置组提供的物理存储区域上形成至少一个以上逻辑卷，把该逻辑卷提供给主机（关于更详细的内容参见“在主机上运行的数据库程序”）。主机可以通过发送规定的指令对逻辑卷进行写入、读出。

随着信息化社会的发展，应该用数据库管理的数据与日剧增。因此，要求更高性能、更大容量的存储控制装置，为了响应这一市场需求，新型的存储控制装置正在不断开发中。作为将存储系统导入新型存储控制装置的方法，设想了两种方法。其一是将旧型号的存储控制装置和新型号的存储控制装置完全替换，用全新的存储控制装置构成存储系统（专利文献 1）的方法。另一种是，在由旧型号的存储控制装置组成的存储系统上新追加新型号的存储控制装置，使新旧存储控制装置并存的方法。

【专利文献 1】特开平 10—508967 号公报

在从旧型号的存储控制装置完全转换到新型号的存储控制装置的场合（专利文献 1），可以利用新型号的存储控制装置的功能、性能，但是，不能有效利用旧型号的存储控制装置。另一方面，在使旧型号的存储控制装置和新型号的存储控制装置并存的场合，构成存储系统的存储控制装置的数目增大，管理、运用新旧存储控制装置的时间、劳力大。

发明内容

本发明是鉴于上述问题而做出的，本发明的一个目的是提供一种存储系统以及存储控制装置，该存储系统以及存储控制装置能够防止负荷

集中在特定的存储控制装置上并使负荷分散。

本发明的另一个目的是提供一种存储系统以及存储控制装置，该存储系统以及存储控制装置能够使如新旧存储控制装置这样不同的存储控制装置协作，同时实现存储资源的逻辑的集中化和负荷的分散。

本发明的再一个目的是提供一种存储系统以及存储控制装置，该存储系统以及存储控制装置能够在实现例如存储系统的高性能、大容量的同时，防止处理负荷集中在高功能高性能的存储控制装置上，使负荷分散。

本发明的其它目的根据后面对实施例的说明可以明了。

为了解决上述问题，依据本发明的存储系统是连接第一存储控制装置和第二存储控制装置使其能够通信而构成，进行响应来自上位装置的请求的数据处理的存储系统，其特征在于，第一存储控制装置具有第一控制装置，该第一控制装置判断第二存储控制装置能否执行涉及从上位装置接收的第一请求的特定的数据处理，在判断为第二存储控制装置能够执行的场合，生成与第一请求对应的第二请求，并发送给第二存储控制装置；第二存储控制装置具有第二控制装置，该第二控制装置根据从第一存储控制装置接收的第二请求，进行特定的数据处理。

作为存储控制装置，可以举出例如磁盘阵列装置或者纤维通道开关等。作为上位装置，可以举出例如个人计算机、主机等计算机。第一存储控制装置和第二存储控制装置通过通信网络连接起来可以进行双向通信，第一存储控制装置和上位装置也通过通信网络连接起来可以进行双向通信。另外，第二存储控制装置和上位装置之间也连接起来可以进行双向通信。作为通信网络，可以举出例如 LAN（Local Area Network）、SAN（Storage Area Network）、专用线路、因特网等。

上位装置向第一存储控制装置发送第一请求。该请求包含例如用于特定要求内容的指令码、特定作为对象的数据的地址的地址信息等而构成。第一存储控制装置的第一控制装置接收第一请求后，判断在第二存储控制装置能否执行在第一请求中请求的特定的处理。这里，作为特定的处理，可以举出各种数据处理，但是，也可以举出单纯的数据输入输

出以外的支持功能（附加的功能）。具体地讲，例如，将如数据的备份、双卷（pair volume）之间的复制、镜像等处理负荷比较大的处理作为特定的处理的话，可以使大的负荷分散到第一存储控制装置以外的装置上。

第一控制装置在判定为第二存储控制装置能够执行特定的处理的场合，生成与第一请求对应的第二请求并发送给第二存储控制装置。第二请求是请求第二存储控制装置进行特定处理的信息。第二存储控制装置接收第二请求后，第二控制装置根据第二请求，执行请求的特定的处理。这样，通过第一请求要求第一存储控制装置进行的数据处理，可以通过第二请求使第二存储控制装置代替执行。

因此，可以使第一存储控制装置的处理负荷分散到第二存储控制装置，防止过大的负荷集中在第一存储控制装置上。于是，可以将第一存储控制装置的信息处理资源（CPU 处理能力或者存储器容量等）分配给向上位装置提供服务。在第一存储控制装置是新型的存储控制装置的场合，可以把第一存储控制装置具有的高性能服务有效地提供给上位装置，可以提高存储系统整体的效率。

在本发明的一个实施例中，第一存储控制装置把第二存储控制装置管理的第二存储区域作为自己管理的第一存储区域假想地提供给上位装置，第一请求是请求涉及第一存储区域的数据处理的请求。

第二存储控制装置具备例如通过磁盘驱动器等存储装置提供的现实的存储区域。在本说明书中将这种实际存在的存储区域称为实存储区域。第一存储控制装置把为第二存储控制装置的实存储区域的第二存储区域当作如同自己有的存储区域，向上位装置假想地提供。因此，第一存储控制装置自身没有必要具备物理的存储区域，可以将具有微计算机系统的智能化的开关机构（纤维通道开关等）作为第一存储控制装置使用。另外，第一存储控制装置也可以是具备物理的存储区域的磁盘阵列装置（磁盘阵列子系统）等。在以磁盘阵列装置构成第一存储控制装置的场合，收进第二存储控制装置的存储区域的结果，可以对上位装置提供比实际具有的存储容量大的存储区域。

上位装置通过发布第一请求，对假想提供的第一存储区域请求数据

操作，但是，实际的数据存储在第二存储控制装置的第二存储区域。由于第一请求以其实体存在于第二存储区域的第一存储区域为对象，所以可以使第二存储控制装置执行请求的特定的数据处理。在来自上位装置的请求是例如第一存储控制装置的状态请求或者控制信息的备份处理等应该在第一存储控制装置自身处理（或者处理合适）的数据处理的场合，可以不委托第二存储控制装置，由第一存储控制装置自身处理，响应上位装置。

在本发明的一个实施例中，第一存储控制装置保持表示第一存储区域和第二存储区域的对应关系的存储区域对应信息，根据该存储区域对应信息向上位装置假想地提供第一存储区域；第一控制装置能够根据存储区域对应信息，执行根据第一请求进行的以第一存储区域为对象的数据处理。

存储区域对应信息是表示作为第一存储区域分配的第二存储区域的对应关系的信息，例如，作为映射表等保持在第一存储控制装置内的存储装置（半导体存储器等）中。存储区域对应信息，例如在定义存储系统的结构等时，可以通过操作员的手动操作、或者通过自动的处理而生成并存储。

在本发明的一个实施例中，第二请求具有与第一请求同样的数据结构而构成。

通过使第一请求和第二请求为同样的数据结构，接收第二请求的第二存储控制装置可以与从上位装置直接指令同样地进行数据处理。也就是说，如果使第二请求采用与第一请求不同的数据结构的话，必须向第二存储控制装置追加用于接收并解释第二请求的功能，但是，通过使第一请求和第二请求为同样的数据结构，无需向第二存储控制装置追加特别的功能，可以有效利用第二存储控制装置。

在本发明的一个实施例中，第一控制装置在向第二存储控制装置发送第二请求前，确认第二存储控制装置能否执行第二请求所涉及的特定的数据处理。

通过在发送第二请求前，确认第二存储控制装置能否执行第二请求

要求的特定的数据处理，可以事前防止向第二存储控制装置发送无用的请求。另外，在第二存储控制装置不能执行特定的处理、第一存储控制装置执行的场合，可以不用发送无用的请求等待从第二存储控制装置返回错误响应，而可以立即执行特定的处理。

在本发明的一个实施例中，第一存储控制装置保持表示可以在第二存储控制装置执行的数据处理功能的功能管理信息，第一控制装置根据功能管理信息判断能否在第二存储控制装置执行第二请求所涉及的特定的数据处理。

功能管理信息可以通过例如按数据的备份、双卷之间的复制、镜像等各种功能，使上述功能与表示能否（可否利用）实行该功能的信息对应而构成。或者，也可以通过功能管理信息只管理在第二存储控制装置中可以实行的功能。

那么，功能管理信息可以在定义存储系统的结构时手动或者自动生成。

在本发明的一个实施例中，具备与第一存储控制装置以及第二存储控制装置两者连接并能够进行通信的备份装置，第一控制装置，在第一请求所涉及的数据处理是要把存储在第一存储区域中的信息转发到备份装置并存储的备份处理的场合，判断第二存储控制装置能否执行备份处理，在判断为第二存储控制装置能够执行备份处理的场合，通过将包含在第一请求中的第一存储区域中的地址变换为第二存储区域中的地址，生成第二请求并发送给第二存储控制装置；第二控制装置根据第二请求，把存储在第二存储区域中的信息转发到备份装置并存储。

在存储系统具备用于取得数据的备份的备份装置的场合，上位装置定期或不定期地把请求备份处理的第一请求发送给第一存储控制装置。接收第一请求的第一存储控制装置（第一控制装置）判断第二存储控制装置能否实行备份处理，在判断为能够执行的场合，生成第二请求，并发送给第二存储控制装置。接收第二请求的第二存储控制装置，如同从上位装置接受了直接指令，把存储在第二存储区域中的信息（数据或者控制信息）转发给备份装置并存储。因此，第一存储控制装置可以使自

身接受的备份请求转嫁给第二存储控制装置，这样，可以把自身的信息处理资源使用在向上位装置提供其它服务上。

在本发明的一个实施例中，第一存储控制装置是进一步假想地提供和第一存储区域成对的第一辅助存储区域的装置，第二存储控制装置进一步具有和第二存储区域成对的第二辅助存储区域，第一控制装置在第一请求所涉及的特定的数据处理是把存储在第一存储区域中的信息复制到第一辅助存储区域的内部复制处理的场合，判断第二存储控制装置能否执行内部复制处理，在判断为第二存储控制装置能够执行内部复制处理的场合，通过将包含在第一请求中的第一存储区域中的地址变换为第二存储区域中的地址，生成第二请求，并发送给第二存储控制装置；第二控制装置根据第二请求，把存储在第二存储区域中的信息复制到第二辅助存储区域。

第一存储控制装置以及第二存储控制装置分别具备两个存储区域。一个存储区域是主存储区域，另一个存储区域是辅助存储区域。主存储区域和辅助存储区域成对，分别存储同一数据。在上位装置对第一存储控制装置请求使存储在第一主存储区域中的信息复制到第一辅助存储区域中的内部复制处理的场合，第一控制装置判断第二存储控制装置能否执行该内部复制处理，在判断为能够执行的场合，把第一请求中的地址变换成第二存储控制装置用的地址，生成第二请求。即，由于第一请求以第一存储控制装置的第一存储区域为对象，因此包含在第一请求中的地址成为表示第一存储区域中的特定存储空间的地址。于是，第一控制装置通过把第一存储区域中的地址变换成第二存储区域中的对应的地址，生成第二请求。由此，第二存储控制装置如同自身直接从上位装置接收指令的场合那样，把存储在为实存储区域的第二主存储区域中的信息存储在为实存储区域的第二辅助存储区域中。因此，第一存储控制装置可以使第二存储控制装置代行内部复制处理，把自身的信息处理资源使用在其它服务上。

在本发明的一个实施例中，和设置第一存储控制装置以及第二存储控制装置设置的主场所成对，有副场所和主场所连接并可以进行通信，

该副场所具备另外的第一存储控制装置以及另外的第二存储控制装置，该另外的第一存储控制装置把另外的第二存储控制装置管理的另外的第二存储区域作为自己管理的另外的第一存储区域假想地提供，主场所的第一控制装置在第一请求所涉及的数据处理是要把存储在第一存储区域中的信息复制到副场所的另外的第一存储区域的外部复制处理的场合，判断第二存储控制装置以及另外的第二存储控制装置双方能否执行外部复制处理，在判断为各第二存储控制装置分别能够执行外部复制处理的场合，生成与第一请求对应的第二请求，并发送给第二存储控制装置，第二控制装置根据第二请求，通过使存储在第二存储区域中的信息复制到另外的第二存储区域中，来执行外部复制处理。

存储系统可以由主场所（主要场所，本地场所）、和设置在离开主场所的地方的副场所（辅助场所，远离场所）双方构成。主场所具备上位装置、第一存储控制装置和第二存储控制装置，副场所具备另外的第一存储控制装置以及另外的第二存储控制装置。副场所是主场所的备份用场所，主场所的第一存储控制装置和副场所的另外的第一存储控制装置构成对。于是，在正副各场所，第一存储控制装置（另外的第一存储控制装置）把第二存储控制装置（另外的第二存储控制装置）现实提供的实际的存储区域（第二存储区域，另外的第二存储区域）当作是自身的存储区域（第一存储区域，另外的第一存储区域）。因此，正副各场所的第二存储控制装置和另外的第二存储控制装置互相构成对。

设置在主场所的第一存储控制装置假想地提供给上位装置的第一存储区域是主卷，设置在副场所的第一存储控制装置假想地提供给上位装置的第一存储区域是和主卷成对的辅助卷。上位装置定期或者不定期地向主场所的第一存储控制装置请求外部复制处理。所谓外部复制处理是使主场所的信息复制到副场所的处理。通过第一请求要求外部复制处理后，属于主场所的第一控制装置判断正副两场所的第二存储控制装置能否执行外部复制处理。在分别设置在两场所的第二存储控制装置能够执行外部复制处理的场合，生成与第一请求对应的第二请求，并将该第二请求发送给设置在主场所的第二存储控制装置。由此，设置在主场所的

第二存储控制装置读出存储在第二存储区域中的数据，并发送给设置在副场所的第二存储控制装置，使其存储在副场所的第二存储区域中。这里，通过把在第一请求中明示的复制源地址和复制目的地地址变成正副两场所的第二存储区域中的地址，来生成第二请求。由此，实际上在存储数据或控制信息的正副两场所的第二存储控制装置之间进行外部复制处理，正副两场所的第一存储控制装置不直接参与外部数据处理。因此，设置在主场所的第一存储控制装置不为执行外部复制处理消费信息处理资源，可以将节省的信息处理资源用在向上位装置提供服务上。

这里，位于副场所的另外的第一存储控制装置保持表示在另外的第二存储控制装置可以执行的数据处理功能的另外的功能管理信息，主场所的第一控制装置在把第二请求发送给第二存储控制装置之前，通过向副场所的另外的第一存储控制装置问询，可以判断另外的第二存储控制装置能否执行外部复制处理。

也就是说，设置在副场所的另外的第二存储控制装置可以执行的数据处理功能的信息作为副场所用的功能管理信息，由设置在副场所的另外的第一存储控制装置保持。因此，设置在主场所的第一存储控制装置在发送第二请求之前，向设置在副场所的另外的第一存储控制装置问询，判断副场所的另外的第二存储控制装置能否执行外部复制处理。此外，在设置在主场所的第二存储控制装置能够执行的数据处理功能的信息，作为主场所用的功能管理信息，由设置在主场所的第一存储控制装置保持。

在本发明的一个实施例中，第一存储控制装置保持表示第一存储区域和第二存储区域的对应关系的存储区域对应信息，另外的第一存储控制装置保持表示另外的第一存储区域和另外的第二存储区域的对应关系的另外的存储区域对应信息，第一控制装置在向第二存储控制装置发送第二请求的场合，发送各存储区域对应信息。

由此，设置在主场所的第二存储控制装置可以把握涉及复制源以及复制目的地的存储空间的信息，进行外部复制处理。

在本发明的一个实施例中，第一存储控制装置具备保持涉及在外部

复制处理中被上位装置更新过的第一存储区域的信息的信息的更新位置信息保持装置，第一控制装置在外部复制处理结束的场合，为了把在第一存储区域中更新过的信息存储在另外的第一存储区域中，根据更新位置信息保持装置，生成第二请求，从第二存储控制装置读出更新过的信息，并发送读出的信息。

亦即，在分别设置在正副两场所的第二存储控制装置之间进行外部复制处理期间，有时上位装置访问设置在主场所的第一存储控制装置并使数据更新。因此，将涉及在外部复制处理中更新过的信息的信息（例如，更新过的逻辑块地址的信息等）存储在更新位置信息保持装置中，在外部复制处理结束后，使更新过的信息反映在副场所的第二存储区域中。

遵从本发明的另外的观点的控制方法是用于控制存储系统的控制方法，所述存储系统通过连接第一存储控制装置和第二存储控制装置并使其能够进行通信而构成，进行响应来自上位装置的请求的数据处理；其中，第一存储控制装置执行：从上位装置接收第一请求的步骤；判断第二存储控制装置能否执行接受的第一请求所涉及的特定的数据处理的步骤；在判断为第二存储控制装置能够执行的场合，生成与第一请求对应的第二请求的步骤；把生成的第二请求发送给第二存储控制装置的步骤；第二存储控制装置执行：从第一存储控制装置接收第二请求的步骤；根据接收的第二请求进行特定的数据处理的步骤。

遵从本发明的进一步另外的观点的存储控制装置是连接第二存储控制装置以及上位装置并使其能够进行通信，进行响应来自上位装置的请求的数据处理的存储控制装置，其中，该存储控制装置由：接收来自上位装置的请求的接收装置；判断第二存储控制装置能否执行接收的第一请求所涉及的特定的数据处理的判断装置；在判断为第二存储控制装置能够执行的场合，生成与第一请求对应的第二请求的请求装置；和将生成的第二请求发送给第二存储控制装置的发送装置而构成。

遵从本发明的程序是控制与第二存储控制装置以及上位装置连接并能够与其进行通信、进行响应来自上位装置的请求的数据处理的第一存

储控制装置的程序；其中，该程序在第一存储控制装置的计算机上实现判断第二存储控制装置能否执行从上位装置接收的第一请求所涉及的特定的数据处理的功能、在判断为第二存储控制装置能够执行的场合生成与第一请求对应的第二请求的功能、和将生成的第二请求发送给第二存储控制装置的功能。

附图说明

图 1 是概括地表示根据本发明的第一实施例的存储系统的主要部分的框图。

图 2 是由磁盘阵列装置构成第一存储控制装置的场合的框图。

图 3 是表示从主机装置看图 2 所示的磁盘阵列装置的场合的逻辑概况结构的说明图。

图 4 是将第一存储控制装置作为纤维通道开关而构成的场合的框图。

图 5 是表示从主机装置看图 4 所示的纤维通道开关的场合的逻辑概况结构的说明图。

图 6 是表示从第一存储控制装置对第二存储控制装置读写数据的场合的一例的概况框图。

图 7 是作为磁盘阵列装置构成第一存储控制装置，与一台第二存储控制装置连接的场合的概况框图。

图 8 是作为纤维通道开关构成第一存储控制装置，由两个实 LU 构成假想的 LU 的场合的概况框图。

图 9（a）表示作为纤维通道开关构成第一存储控制装置的场合的映射表，图 9（b）表示作为磁盘阵列装置构成第一存储控制装置的场合的映射表，图 9（c）表示由多个实 LU 构成假想的 LU 的场合的映射表。

图 10 是第一存储控制装置自己进行直接备份的场合的说明图。

图 11 是具有假想的 LU 的第一存储控制装置进行直接备份的场合的说明图。

图 12 是具有假想的 LU 的第一存储控制装置使第二存储控制装置代行直接备份的场合的说明图。

图 13（a）表示功能管理表，图 13（b）表示从主装置向第一存储控

制装置发送的指令的数据结构，图 13（c）表示从第一存储控制装置向第二存储控制装置发送的指令的数据结构。

图 14 是表示在进行数据的直接备份的场合的第一存储控制装置中的处理概要的流程图。

图 15 是概括地表示涉及本发明的第二实施例的存储系统的整体的框图。

图 16（a）表示主场所侧的映射表，图 16（b）表示副场所侧的映射表，图 16（c）表示初始复制开始指令的数据结构。

图 17 是表示在进行初始复制的场合的第一存储控制装置中的处理的概要的流程图。

图 18 是表示具备假想的 LU 的第一存储控制装置使第二存储控制装置进行逻辑卷复制的场合的存储系统的全体概要的框图。

图 19（a）表示主场所侧映射表，图 19（b）表示副场所侧映射表，图 19（c）表示主场所侧功能管理表，图 19（d）表示副场所侧功能管理表，图 19（e）表示初始复制开始指令的数据结构。

图 20 是表示通过第一存储控制装置执行的初始复制处理的概要的流程图。

图 21 是表示进行从第一存储控制装置依赖的初始复制的第二存储控制装置的处理概要的流程图。

图 22 涉及本发明的第三实施例，是表示进行内部卷的同步化的场合的结构概要的框图。

具体实施方式

下面根据图 1～图 22 说明本发明的实施例。

在本发明中，如以下详细说明的，第一存储控制装置把第二存储控制装置提供的实存储区域当作自身的存储区域假想地提供给主机，同时对于在第二存储控制装置能够执行的处理委托第二存储控制装置进行。亦即，在本发明中，可以将分散的物理存储资源逻辑上集中在第一存储控制装置进行管理，进一步，在能够集中管理的状态下，分散执行负荷大的特定处理。

这里，第一存储控制装置可以具备多种动作模式。第一模式是本发明特有的模式，是在对主机提供假想的存储区域的同时，对于在第二存储控制装置能够执行的处理委托第二存储控制装置进行的模式。第二模式是本发明的前提构成中的模式，是在对主机提供假想的存储区域的同时，通过第一存储控制装置处理在第二存储控制装置能够执行的处理的模式。第三模式是现有技术已知的模式，是不提供假想的存储区域，第一存储控制装置对于实际的存储区域进行处理的模式。依据本发明的存储系统至少具备第一模式。进一步，也可以具有第二模式、第三模式。具有多种动作模式的存储系统不是现有技术中公知的技术，是本发明的特征之一。

【第一实施例】

[整体结构的概要]

首先，根据图 1～图 9，说明依据本发明的实施例的构成。图 1 是表示根据本实施例的存储系统的主要部分的构成的框图。

主装置 1 是具备例如 CPU（Central Processing Unit）或存储器等信息处理资源的计算机装置，例如，作为个人计算机、工作站、主机等构成。主装置 1 具有例如键盘开关、或者指点装置、微音器等信息输入装置（未图示）和例如监视器或者扬声器等信息输出装置（未图示）。在主装置 1 上，安装有用于管理 RAID 的 RAID 管理器或者数据库管理程序等。

主装置 1 通过通信网络 2 分别与后面要说明的第一存储控制装置 10 以及第二存储控制装置 200 连接，并且可以进行双向通信。作为通信网络 2 可以根据场合适当选用例如 LAN、SAN、因特网、专用线路、公共线路等。通过 LAN 的数据通信，例如可以遵照 TCP/IP（Transmission Control Protocol/Internet Protocol）协议进行。在主装置 1 通过 LAN 与第一存储控制装置 10 等连接的场合，主装置 1 指定文件名请求以文件为单位的数据输入输出。在主装置 1 通过 SAN 与第一存储控制装置 10 等连接的场合，主装置 1 遵照纤维通道协议以作为通过多个磁盘存储装置（磁盘驱动器）提供的存储区域的数据管理单位的块为单位，请求数据输入输出。

此外，主装置 1 和第一存储控制装置 10 之间、第一存储控制装置 10 和第二存储控制装置 200 之间、主装置 1 和第二存储控制装置 200 之间也可以分别通过不同的通信网络连接，也可以通过如图所示共同的通信网络 12 连接。在第一存储控制装置 10 和主装置 1 之间的通信网络与第一存储控制装置 10 和第二存储控制装置 200 之间的通信网络不同的场合，进行协议变换等从第一存储控制装置 10 向第二存储控制装置 200 发送第二请求的话即可。

第一存储控制装置 10 是如后面要说明的作为例如磁盘阵列装置或者纤维通道开关而构成的计算机系统。第一存储控制装置 10 把第二存储控制装置 200 提供的在物理存储区域上设定的逻辑存储区域（逻辑卷（Logical Unit））当作第一存储控制装置自身提供的，假想地提供给主装置 1。

在下面的说明中，有时把逻辑卷称为 LU，把第一存储控制装置 10 假想地提供的 LU 称为假想的 LU。另外，有时把现实中存在的 LU 称为实 LU。在图中，以假想线（双点划线）表示假想的 LU，以实线表示实 LU。此外，在由磁盘阵列装置构成第一存储控制装置 10 的场合，第一存储控制装置 10 具有假想的 LU 和实 LU 两者。第一存储控制装置 10 内的实 LU 可以理解为其物理的实存储区域在第一存储控制装置 10 内或者是在第一存储控制装置 10 的直接管理下的假想的 LU。这样考虑的话，也可以把存储区域的实体位于第二存储控制装置 200 侧的假想的 LU 称为假想的外部 LU，反之，把存储区域的实体位于第一存储控制装置 10 侧的假想的 LU 称为假想的内部 LU。

这样，由于第一存储控制装置 10 把第二存储控制装置 200 的 LU 作为假想的 LU 提供给主装置 1，所以第一存储控制装置 10 自身不必具备实 LU。因此，即使是不具备提供物理存储区域的磁盘驱动器的纤维通道开关，只要具有必要的信息处理能力，就可以作为第一存储控制装置 10 使用。

第一存储控制装置 10 具备用于在主装置 1 或者第二存储控制装置 200 之间进行通信的端口 11、12 和指令控制单元 210。另外，在由磁盘

阵列装置构成第一存储控制装置 10 的场合，第一存储控制装置 10 具备存储装置 70。存储装置 70 具备多个例如硬盘、软盘、磁带、半导体存储器、光盘等装置而构成，在这些物理的存储区域上设定 LU71。另外，第一存储控制装置 10 提供映射第二存储控制装置 200 的实 LU 而成的假想的 LU72。在图 1 中，为了便于说明，在存储装置 70 内包含实 LU71 以及假想的 LU72。存储装置 70 可以直接和第一存储控制装置 10 连接，或者，也可以通过通信网络连接。另外，也可以将存储装置 70 和第一存储控制装置 10 一体化。

各端口 11、12 是进行数据的收发的装置。一方的端口 11 是接收来自主装置 1 的请求的目标端口（TP），另一方的端口 12 是向第二存储控制装置 200 发送请求的启动端口（IP）。两端口 11、12 的构造相同，根据数据通信上的作用，成为目标端口或者启动端口。

指令控制单元 100 通过第一存储控制装置 10 具有的信息处理资源（CPU，存储器，输入输出电路等）或者软件而实现。指令控制单元 100 可以通过例如第一存储控制装置 10 的主控制器（未图示）而实现，也可以通过通道适配器或者磁盘适配器的协作而实现。指令控制单元 100 具备指令解析单元 110、判断单元 120、指令生成单元 130、指令表 TC、映射表 TM 以及功能管理表 TF。

指令解析单元 110 是根据指令表 TC 解析通过端口 11 从主装置 1 接收的指令（请求）的装置。在指令表 TC 中，预先登录有各种指令，可以通过参照指令表 TC，判别包含在来自主装置 1 的请求中的指令代码是请求什么的指令。

判断单元 120 是判断在第二存储控制装置 200 能否执行从主装置 1 接收的请求所涉及的数据处理的装置。判断单元 120 通过参照功能管理表 TF，判定第二存储控制装置 200 是否具备所请求的数据处理功能。功能管理表 TF 的详细情况后面说明，不过，功能管理表 TF 是在例如定义存储系统的构成等时生成的，所以，登录有安装在第二存储控制装置 200 中的支持功能。

指令生成单元 130 是将从主装置 1 接收的请求变换为第二存储控制

装置200用的请求的装置。指令生成单元130通过参照指令表TC以及映射表TM，生成第二存储控制装置200能够执行主装置1请求的数据处理的请求。映射表TM是表示第二存储控制装置200的LU和第一存储控制装置10的假想的LU的对应关系的表。指令生成单元130通过参照映射表TM，把以假想的LU的存储空间为对象的地址变换为以实LU的存储空间为对象的地址。

这样，指令控制单元100解析从主装置1接收的第一请求，判断第二存储控制装置200能否执行通过第一请求请求的数据处理(支持功能)。然后，在判断为第二存储控制装置200能够执行的场合，指令控制单元100生成与第一请求对应的第二请求，把该第二请求通过通信网络2从端口12发送给第二存储控制装置200。在第二存储控制装置200不能执行从主装置1请求的数据处理的场合，第一存储控制装置10执行请求的数据处理。

[将第一存储控制装置适用于磁盘阵列装置的场合]

图2是表示作为磁盘阵列装置构成第一存储控制装置10的场合的具体例子的框图。第一存储控制装置10如后面分别说明的，可以由例如多个通道适配器20、超高速缓冲存储器40、公共存储器50、多个磁盘适配器60、存储装置70等构成。另外，虽然图中作了省略，但是第一存储控制装置10可以具备例如用于控制第一存储控制装置10的整体动作的MPU（Micro Processing Unit），和用于执行环境设定或者管理各种状态等的维护管理用终端。

通道适配器（CHA）20是进行与主装置1之间的数据通信的装置。各通道适配器20具备用于进行与主装置1等通信的通信端口21、用于转发接收的数据的转发单元22、控制通道适配器内的动作的微处理器（略记为MP）23、和存储器24。通过微处理器23实现用于解释并处理从主装置1接收的各种指令的指令控制单元100。或者，也可以通过总括装置整体的MPU实现指令控制单元100，再有，也可以通过MPU和通道适配器20的协作、MPU和通道适配器20以及磁盘适配器60的协作、通道适配器20和磁盘适配器60的协作，实现指令控制单元100。

给各通道适配器 20 分配用于识别各个通道适配器的网络地址（例如 IP 地址或者 WWN（World Wide Name）），各通道适配器 20 分别作为单个的 DAS（Direct Attached Storage）或者作为 NAS（Network Attached Storage）行动。也就是说，各通道适配器 20 可以分别个别地接受来自各主装置 1 的请求。各通道适配器 20 可以具备多个由端口 21、转发单元 22、微处理器 23 以及存储器 24 构成的控制电路而构成。

各微处理器 23 通过主侧公共存储器存取电路 31 分别与公共存储器 50 连接，向公共存储器 50 写入控制信息，或者参照写入到公共存储器 50 中的控制信息。各转发单元 22 通过主侧超高速缓冲存储器存取电路 32 分别与超高速缓冲存储器 40 连接，使从主装置 1 等接收的数据存储在超高速缓冲存储器 40 中，另外，读出存储在超高速缓冲存储器 40 的数据，并发送给主装置 1。超高速缓冲存储器 40 以及公共存储器 50 是由各通道适配器 20 和各磁盘适配器 60 共有的存储装置。在超高速缓冲存储器 40 中主要存储数据，在公共存储器 50 中主要存储控制信息或者指令等。另外，在公共存储器 50 中还设定有工作区域。上述指令表 TC、映射表 TM、功能管理表 TF 可以存储在例如公共存储器 50 中，或者也可以存储在超高速缓冲存储器 40 中。

各磁盘适配器（DKA）60 是管理与存储装置 70 的磁盘装置组 73 之间的数据输入输出的装置。磁盘适配器 60 把通道适配器 20 从主装置 1 接收的数据，按照来自主装置 1 的写入要求，写入磁盘装置组 73 的特定地址。此时，各磁盘适配器 60 把逻辑卷中的逻辑地址变换成物理磁盘中的物理地址。磁盘适配器 60 在通过 RAID 管理磁盘装置组 73 的场合，进行依照 RAID 构成的数据存取，也可以进行存储在磁盘装置组 73 中的数据的复制管理控制或者备份控制。再有，磁盘适配器 60 也可以以防止在灾害发生时的数据消失等为目的，进行把主场所的数据的复制品存储在副场所的控制（复制功能或者远程复制功能）等。

各磁盘适配器 60 可以分别具备多个由通信端口 61、转发单元 62、微处理器 63、存储器 64 组成的控制电路。各通信端口 61 进行存储装置 70 的磁盘装置组 73 之间的数据通信。各转发单元 62 通过装置侧超高速

缓冲存储器存取电路 34 与超高速缓冲存储器 40 连接，把写入超高速缓冲存储器 40 中的数据转发给磁盘装置组 73，或者把从磁盘装置组 73 读出的数据写入到超高速缓冲存储器 40。微处理器 63 通过装置侧公共存储器存取电路 33 与公共存储器 50 连接，可以参照写入公共存储器 50 的控制信息或者指令。

主侧公共存储器存取电路 31 以及装置侧公共存储器存取电路 33，和主侧超高速缓冲存储器存取电路 32 以及装置侧超高速缓冲存储器存取电路 34 可以作为例如通过高速开关动作进行数据传送的超高速交叉开关等那样的高速总线构成。

图 3 是表示从主装置 1 看第一存储控制装置 10 的场合的逻辑结构的主要部分的框图。第一存储控制装置 10 对主装置 1 提供两种 LU。其一是通过第一存储控制装置 10 直接管理的存储装置 70 的物理磁盘装置组 73 提供的实 LU71。另一是把通过第二存储控制装置 200 管理的存储装置 220 的磁盘装置组 221 提供的实 LU 作为第一存储控制装置 10 的 LU 提供的假想的 LU72。如图 3 所示，实 LU71 以及假想的 LU72 中任何一个都可以设置多个。各 LU71、71 分别由多个磁盘装置组构成。

[将第一存储控制装置适用于纤维通道开关的场合]

下面，图 4 是表示把第一存储控制装置 10 作为智能化的纤维通道开关而构成的场合的具体例子的框图。为了与作为磁盘阵列装置构成的场合区别，在图中，在符号 10 上附加（SW）。第一存储控制装置 10 具备多个通道适配器 20、公共存储器存取电路 31、33、超高速缓冲存储器存取电路 32、34、高速缓冲存储器 40、公共存储器 50、和控制器 80 而构成。通道适配器 20 等的详细情况因为与图 2 所述相同，因此省略。控制器 80 是总括并控制整体的动作的装置，具备 MPU 或者存储器。和图 2 所示的装置的大的不同点是作为纤维通道开关而构成的第一存储控制装置 10 不具备直接管理的存储装置 70 这一点。

图 5 是表示从主装置 1 看作为纤维通道开关而构成的第一存储控制装置 10 时的逻辑结构的主要部分的框图。如图所示，第一存储控制装置 10 没有实 LU，但是具备至少一个以上的假想的 LU72。如以上所述，假

想的 LU72 的实体存在于第二存储控制装置 200 的存储装置 220 中。

[向假想的 LU 的数据存取]

下面参考图 6，说明向假想的 LU72 的数据存取。图 6 表示作为纤维通道开关构成第一存储控制装置 10 的场合。

在主装置 1 对假想的 LU72 请求写入数据或者读出数据的场合，主装置 1 向第一存储控制装置 10 发布指令 CH。在该指令 CH 中包含有用于特定第一存储控制装置 10 的信息（端口 ID，WWN 等）、表示指令的种类的指令代码（写指令，读指令等）、作为对象的数据的读出地址（在读指令的场合）等。来自主装置 1 的指令 CH 通过 SAN 等通信网络 2，在第一存储控制装置 10 的目标端口 11 接收，输入到指令控制单元 100。

指令控制单元 100 解析接受的指令 CH，通过参照映射表 TM，进行从主装置 1 请求的数据处理。在由纤维通道开关构成第一存储控制装置 10 的场合，映射表 TM 例如可以如图 9（a）所示地构成。映射表 TM 例如可以通过使用于识别假想的 LU（假想的逻辑卷）72 的卷 ID（Vol ID）、设定在假想的 LU72 中的逻辑块地址（BLK ADDR）、用于识别具有与假想的 LU72 对应的实 LU222 的第二存储控制装置 200 的装置 ID、用于识别与实 LU222 对应的端口的端口 ID、用于识别实 LU222 的卷 ID（Vol ID）、与假想的 LU72 的逻辑块地址对应设定在实 LU222 中的逻辑块地址分别对应而构成。因此，通过参照映射表 TM，可以掌握假想的 LU72 的特定的逻辑块地址与哪一个第二存储控制装置 200 提供的实 LU222 的哪个逻辑块地址（以下简记为“地址”）对应。该映射表 TM 可以在例如构成存储系统、登录逻辑卷时，手动或者自动地登录。

这里，作为假想 LU 与实 LU 之间的映射，除了在对于全部地址使用映射表的场合之外，有时在参照表之外通过一些计算求对应的地址。在后者的方法中，与前者相比可以减小映射表的大小。在本发明中的映射中，通过参照表及计算，可以求对应的地址。

这样，主装置 1 发布用于使数据写入假想的 LU72 的地址 Bb 到 Bc 的范围内的指令 CH。指令控制单元 100 接收指令 CH 后，根据指定的地址参照映射表 TM。由此，指令控制单元 100 在知道假想的 LU72 与通过

可以从端口 201（端口 ID=TP2）存取的卷 ID α 2 特定的实 LU222 对应的同时，掌握假想的 LU72 中的地址 Bb-Bc 与实 LU222 的地址 Bb2-Bc2 对应。因此，指令控制单元 100 通过部分改写从主装置 1 接收的指令 CH 的内容，生成指令 CS。即，指令控制单元 100 通过把包含在从主装置 1 接收的指令 CH 中的卷 ID 和地址根据映射表 TM 改写成实 LU222 的卷 ID 和地址，生成指令 CS（Vol ID α →Vol ID α 2，BLK ADDR Bb-Bc→BLK ADDR Bb2-Bc2）。然后，指令控制单元 100 通过通信网络 2 把生成的指令 CS 从启动端口 12 发送给第二存储控制装置 200。此外，以下的场合也同样，不过，在进行地址变换的场合，不限于仅使用映射表 TM 的场合，可以在参照映射表 TM 之上通过进行一些运算，求对应的地址。

从第一存储控制装置 10 发送的指令 CS 通过第二存储控制装置 200 的目标端口 201 接收，转送给指令控制单元 210。指令控制单元 210 解析指令 CS 的内容，将数据写入物理存在的实 LU222 的指定的地址（Bb2-Bc2）。数据的写入结束后，第二存储控制装置 200 向第一存储控制装置 10 报告写入结束。此外，作为进行向主装置 1 报告写入结束的时间，可以在例如第一存储控制装置 10 接收指令 CH 的时刻（非同步式）、或者在从第二存储控制装置 200 接受写入结束报告的时刻（同步式）的任何一个时刻进行。在同步式的场合，因为产生相当于等待来自第二存储控制装置 200 的响应的时间的延迟，因此适用于第一存储控制装置 10 和第二存储控制装置 200 相离不太远而设置的场合。在第一存储控制装置 10 和第二存储控制装置 200 以远距离相离的场合，因为响应延迟和传播延迟的问题，一般不适用同步式，而采用非同步式。

在主装置 1 从假想的 LU72 读出数据的场合，进行与上述写入数据时同样的处理。主装置 1 发布的指令 CH（读请求）被第一存储控制装置 10 的指令控制单元 100 解析。指令控制单元 100 参照映射表 TM，把作为读出对象指定的假想的 LU72 的地址变换成实 LU222 的地址，生成指令 CS，并把指令 CS 发送给第二存储控制装置 200。指令 CS 通过第二存储控制装置 200 的目标端口 201 接收，被指令控制单元 210 解析。指令控制单元 210 从实 LU222（在已经读出到超高速缓冲存储器的场合从超

高速缓冲存储器）读出指定的地址的数据，并把读出的数据发送给第一存储控制装置 10。第一存储控制装置 10 把接收的数据发送给主装置 1。在从主装置 1 指定的地址的数据已经存储在第一存储控制装置 10 内的超高速缓冲存储器 40 内的场合，也可以把该数据发送给主装置 1。

下面，图 7 表示作为磁盘阵列装置构成第一存储控制装置 10 的场合。在这一场合，第一存储控制装置 10 具备自身直接管理的实 LU71 和假想的 LU72。假设假想的 LU72 的卷 ID 为 α，实 LU71 的卷 ID 为 β。

在第一存储控制装置 10 具备实 LU71 以及假想的 LU72 的场合，映射表 TM 例如可以如图 9（b）所示而构成。在图 9（b）所示的表中，与假想的 LU72 相关的部分和图 9（a）所示的表相同。不同的点是与第一存储控制装置 10 的实 LU71 相关的部分。实 LU71，因为第一存储控制装置 10 自身具有，所以其装置 ID 成为第一存储控制装置 10 的 ID。出于同样的理由，用于访问实 LU71 的端口 ID 中登录有表示是内部逻辑卷的信息（INTERNAL）。另外，因为不存在与第二存储控制装置 200 的 LU 的对应关系，对于对应的卷 ID 或者地址没有登录。通过利用表示端口 ID 字段以及内外的区别的识别符（INTERNAL），可以通过同样构造的映射表 TM 管理实 LU71 和假想的 LU72。此外，不限于此，也可以使用不同的表管理实 LU71 和假想的 LU72。

因为从主装置 1 向假想的 LU72 写入及读出数据与上述纤维通道开关的场合相同，因此省略。

下面，图 8 表示第一存储控制装置 10 作为智能化的纤维通道开关而构成的同时、第一存储控制装置 10 的假想的 LU72 由第二存储控制装置 200A、200B 分别提供的实 LU222A、222B 的两个逻辑卷构成的场合。在这一场合，映射表 TM，例如可以如图 9（c）所示而构成。对于一方的第二存储控制装置 200A，假设其装置 ID 为 SD2（1）、端口 ID 为 TP2（1）、卷 ID 为 α 2A，同样，对于另一方的第二存储控制装置 200B，假设其装置 ID 为 SD2（2）、端口 ID 为 TP2（2）、卷 ID 为 α 2B。

如图 9（c）所示的映射表 TM 所示，可以看到，第一存储控制装置 10 对主装置 1 提供的假想的 LU72（卷 ID= α）由可以从端口 ID=TP2（1）

存取的 LU222A（卷 ID=α 2A）和可以从端口 ID=TP2（2）存取的 LU222B（卷 ID=α 2B）构成。这样，第一存储控制装置 10 可以汇总多个分散的 LU 构筑一个或者多个假想的 LU。因此，例如通过汇总多台使用效率低的旧型号的存储控制装置，构成一台或者多台假想的 LU，提供给主装置 1，可以再次组合存储系统的存储资源，进行有效利用。

[直接备份 1]

下面，说明进行直接备份的场合的流程。所谓直接备份是指不通过主装置 1，在存储控制装置和备份装置之间直接进行数据的备份的处理。直接备份是存储控制装置提供的支持功能的一种。

图 10 是简单地表示可以认为是一般方法的直接备份的构造的说明图。图 10 所示的存储控制装置 10（N）是不构筑假想的 LU 的普通的存储控制装置。备份装置 3 是存储数据全部或其一部分的复制的存储装置。作为备份装置 3，可以采用例如 MO（magneto-optic：光磁型存储装置）、CD-R（CD-Recordable：可读写的小型盘）、DVD-RAM（Digital Versatile Disk-RAM：可读写 DVD）等盘系列存储装置、或者例如 DAT（Digital Audio Tape）磁带、盒式磁带、开放式磁带、卡盘带等磁带系列存储装置等。在图示的例子中，假设是磁带系列存储装置，但是不限于此。

在直接备份中，向备份装置写入数据大致可以区分为两种。其一是数据的备份，另一是控制信息的写入。是进行哪一种备份，由主装置 1 发布的指令（请求）决定。在进行数据备份的场合，从主装置 1 向存储控制装置 10（N）发送如图 10 中的（a）所示的复制请求。在执行数据备份的场合，在从主装置 1 发送的指令 CH 中，分别在指令代码中存储有表示复制请求的请求 ID、在复制源地址中存储有表示进行备份的数据范围的地址以及卷 ID（图中只表示出地址）、在复制目的地装置 ID 中存储有表示用于特定备份装置 3 的装置 ID。另一方面，在进行控制信息的备份的场合，在指令 CH 中，分别在指令代码中存储有表示控制信息的写入请求的请求 ID、在复制目的地装置 ID 中存储有备份装置 3 的 ID，同时在指令 CH 内，包含为复制对象的控制信息。作为这些请求的实例，可以举出例如通过 SCSI-3 规定的 Extended Copy 指令。

首先说明在进行数据备份的场合的动作。主装置 1 生成如图 10（a）所示的指令（复制请求）CH，指定作为 LU71 中的备份对象的逻辑块的地址。生成的指令 CH 通过通信网络 2 从主装置 1 发送给存储控制装置 10（N）。存储控制装置 10（N）通过目标端口 11 接收指令 CH。接收的指令 CH 被指令控制单元 100 解析，识别为请求数据备份的复制请求。指令控制单元 100 通过参照指令 CH 中的复制源地址，读出指定的逻辑卷 β 中指定的地址 Ba-Bd 的逻辑块的数据。指令控制单元 100 生成用于将读出的备份对象的数据写入备份装置 3 的指令（写入请求），并从启动端口 12 向备份装置 3 发送写入指令。从存储控制装置 10（N）接收写入指令的备份装置 3 把接收的数据写入特定的位置，在写入结束的场合，向存储控制装置 10（N）发送写入结束报告。

下面说明进行控制信息的写入的场合的动作。首先，主装置 1 生成如图 10（b）所示的控制信息写入指令。接收来自主装置 1 的指令 CH 的存储控制装置 10（N）解析指令 CH，掌握内容，抽出包含在指令 CH 中的控制信息。然后，和上述同样，指令控制单元 100 生成用于把抽出的控制信息写入备份装置 3 的写入指令，并发送给备份装置 3。从存储控制装置 10（N）接收写入指令的备份装置 3 把控制信息写入规定的位置，并向存储控制装置 10（N）发送写入结束报告。

[直接备份 2]

下面，参照图 11，说明直接备份的另一种方法。在图 11 中，表示把第二存储控制装置 200 的实存储区域当作如同自己的存储区域，向主装置 1 假想地提供，进行直接备份的场合。

如图所示，例如，由智能化的纤维通道开关构成的第一存储控制装置 10、第二存储控制装置 200、主装置 1、和备份装置 3 分别通过通信网络 2 相互连接并可以进行相互通信。于是，如上所述，第二存储控制装置 200 直接管理的实 LU222，通过映射表 TM，分配给第一存储控制装置 10 的假想的 LU72。由此，第一存储控制装置 10 间接地支配第二存储控制装置 200 的实 LU，作为假想的 LU72 提供给主装置 1。

首先说明进行数据的备份的场合的话，首先，主装置 1 生成包含表

示复制请求的请求 ID、用于特定复制对象的数据的卷 ID（α）及地址（Ba-Bd）、用于特定复制目的地装置的装置 ID（备份装置的装置 ID）的指令 CH。接下来，在主装置 1 通过通信网络 2 把指令 CH 发送给第一存储控制装置 10 后，该指令 CH 被第一存储控制装置 10 的目标端口 11 接收。指令控制单元 100 解析指令 CH，通过参照指令 CH 中的复制源地址和映射表 TM，检索与成为复制对象的假想的 LU72 的地址对应的实 LU222 的地址。亦即，指令控制单元 100 在知道与指定直接备份的假想的 LU72 对应的逻辑卷为可从目标端口 201 存取的实 LU222 的同时，知道直接备份的范围为实 LU222 中的地址 Ba2-Bd2 的范围。

接下来，指令控制单元 100 生成用于读出存储在实 LU222 的地址 Ba2-Bd2 中的数据的读出指令（读出请求），从启动端口 12 发送该读出指令。读出指令，通过通信网络 2 在第二存储控制装置 200 的目标端口 201 接收。第二存储控制装置 200 的指令控制单元 210 从实 LU222 读出通过读出指令请求的范围的数据，发送给第一存储控制装置 10。由此，第一存储控制装置 10 取得作为直接备份的对象的数据，并把数据暂时存储在超高速缓冲存储器 40 中。然后，指令控制单元 100 生成用于把从第二存储控制装置 200 取得的数据写入备份装置 3 中的写入指令（复制请求），并把该写入指令向作为复制目的地装置指定的备份装置 3 发送。备份装置 3 根据从第一存储控制装置 10 接收的写入指令，把接收的数据存储在特定的位置。

对进行控制信息的写入的场合进行说明。在这种场合，成为与和图 10 一起说明的被认为是一般的方法的直接备份同样的操作。亦即，主装置 1 生成包含表示控制信息的写入请求的请求 ID、特定复制目的地装置的装置 ID、备份对象的控制信息的指令 CH，并发送给第一存储控制装置 10。指令控制单元 100 解析接受的指令 CH，生成用于将从指令 CH 中抽出的控制信息写入备份装置 3 的写入指令，并发送给备份装置 3。在备份装置 3 接收写入指令后，使控制信息存储在特定的位置。

这样，在把第二存储控制装置 200 的实 LU222 作为第一存储控制装置 10 的存储区域假想地拿进来的构成中，第一存储控制装置 10 主导存

储系统的动作，执行从主装置 1 请求的数据处理（在数据的直接备份的场合）。因此，例如，由于在第一存储控制装置 10 间接支配下的第二存储控制装置 200 的数目、从主装置 1 请求的处理内容、通信网络 2 的速度等诸条件不同而不同，但是，第一存储控制装置 10 的处理负担变重。因此，在本发明中，如以下所述的，通过把在第一存储控制装置 10 能够执行的处理的全部或者一部分使第二存储控制装置 200 执行，来实现负荷分散。

[直接备份 3]

下面，根据图 12～图 14，说明另一个的直接备份的方法。该方法，在可以在第二存储控制装置 200 执行直接备份的场合，使第二存储控制装置 200 执行直接备份这一点上具有特征。

在希望直接备份数据的场合，主装置 1 生成直接备份用的指令 CH。如以上所述，该指令 CH 包含表示复制请求的请求 ID、复制源地址、复制目的地装置的装置 ID。主装置 1 通过通信网络 2 发送指令 CH 后，指令 CH 被第一存储控制装置 10 的目标端口 11 接收。

指令控制单元 100 解析指令 CH 根据请求 ID 并通过参照指令表 TC（省略图示），识别接收的复制请求。指令控制单元 100 根据由指令 CH 指定的复制源地址，参照映射表 TM，检索与假想的 LU72 对应的实 LU222 的地址。亦即，指令控制单元 100 在检测与假想的 LU72 对应的卷的 ID（α2）的同时，检测与假想的 LU72 的存储空间中的地址（Ba-Bd）对应的实 LU222 的存储空间中的地址（Ba2-Bd2）。

接下来，指令控制单元 100 参照功能管理表 TF，确认在第二存储控制装置 200 可以执行的支持功能。图 13（a）表示功能管理表 TF 的一例。图 13（a）表示管理多个第二存储控制装置 200 分别具有的功能的场合。在功能管理表 TF 中，例如使用于分别特定第二存储控制装置 200 的装置 ID（SD2（1）- SD2（n））、用于对各实 LU222 进行存取的端口 ID（TP2（1）- TP2（n））、和能否执行各支持功能（F1 - Fn）的判别信息对应。对于可以执行的功能记录“可”，对于不能执行的功能记录“不可”。作为支持功能，可以举出例如直接备份、双卷的复制、镜像、远程复制等。

该功能管理表 TF，例如在构成存储系统时手动或者自动地登录。在只有一台第二存储控制装置 200 与第一存储控制装置 10 连接的场合，图 13（a）所示功能管理表 TF 的记录成为一个。

在指令控制单元 100 参照功能管理表 TF 确认第二存储控制装置 200 支持直接备份功能后，指令控制单元 100 为了使第二存储控制装置 200 承担直接备份，通过改写第一指令 CH 的一部分，生成第二指令 CS。具体说，例如，如图 13（b）、（c）所示，通过把存储在指令 CH 中的“复制源地址”中的卷 ID 以及地址，参照映射表 TM 分别改写成对应的实 LU222 的卷 ID 以及对应的地址，生成指令 CS。亦即，表示请求的种类的请求 ID 以及复制目的地装置 ID 在两指令 CH、CS 之间通用，只有用于特定作为直接备份的对象的数据的信息变换成实际存储该数据的存储空间的地址。因此，两指令 CH、CS 只有一部分内容不同，数据结构相同。

这样通过指令控制单元 100 组装的指令 CS 从启动端口 12 通过通信网络 2 到达第二存储控制装置 200 的目标端口 201。第二存储控制装置 200 的指令控制单元 210 解析指令 CS，从实 LU222 读出指定的范围的数据。然后，指令控制单元 210 生成用于把该读出的数据写入备份装置 3 的写入指令，并发送给备份装置 3。在备份装置 3 接收来自第二存储控制装置 200 的写入指令后，使接收的数据存储在特定的位置。

参照图 14 来说明第一存储控制装置 10 的动作。在第一存储控制装置 10 从主装置 1 接收指令 CH（S1）后，指令控制单元 100 解析指令 CH（S2），参照功能管理步骤 TF（S3）。然后，在判断为在第二存储控制装置 200 能够执行数据的直接备份的场合（S4：YES），生成用于使第二存储控制装置 200 代行数据的直接备份的指令 CS（S5），并把指令 CS 发送给第二存储控制装置 200（S6）。在判断为第二存储控制装置 200 不支持直接备份功能的场合（S4：NO），指令控制单元 100 通过生成读出指令并发送给第二存储控制装置 200，从第二存储控制装置 200 的实 LU222 中读出备份对象的数据（S7）。指令控制单元 100 通过生成写入指令并发送给备份装置 3（S8），使读出的数据存储在备份装置 3。

此外，在进行控制信息的写入的场合，通过上述的可以认为是一般的方法写入。亦即，指令控制单元 100 从接受的指令 CH 中抽出控制信息，生成用于要把该抽出的控制信息写入备份装置 3 中的写入指令，并发送给备份装置 3。

如以上详细说明的，根据本实施例，判断从主装置 1 请求的数据处理（直接备份）能否在第二存储控制装置 200 执行，在能够在第二存储控制装置 200 执行的场合，可以使第二存储控制装置 200 执行请求的数据处理。因此，由于可以防止处理集中在第一存储控制装置 10 上，使负荷分散，因此可以减轻第一存储控制装置 10 的处理负担。由此，可以为了实现其它的服务向主装置 1 提供第一存储控制装置 10 的信息处理资源，实现存储系统全体的有效利用。

【第二实施例】

[逻辑卷复制 1]

下面根据图 15 说明逻辑卷的复制。根据本实施例的逻辑卷复制在分别离开而设置的主场所（主要场所）和副场所（辅助场所）之间复制逻辑卷的内容。

在主场所，包含主装置 1、由智能化的纤维通道开关构成的第一存储控制装置 10（SW1）、第二存储控制装置 200（1），它们通过通信网络 2（1）相互连接并可以进行相互通信。在副场所，包含由智能化的纤维通道开关构成的第一存储控制装置 10（SW2）、和第二存储控制装置 200（2），它们通过通信网络 2（2）相互连接并可以进行相互通信。主场所的通信网络 2（1）和副场所的通信网络 2（2）也相互连接。以下，在明示为设置在正副哪个场所的装置的场合，追加指主要场所的符号（1）或指辅助场所的符号（2），在没有必要特别区别的场合省略。

根据本实施例的逻辑卷复制是指决定成为复制源的主场所的逻辑卷（α）和成为复制目的地的副场所的逻辑卷（β）的组（双卷），把主场所的逻辑卷（α）的存储内容复制到副场所的逻辑卷（β）上，可以分成两种情况考虑。其一是指把为复制源的主场所的逻辑卷（α）的全部数据发送到为复制目的地的副场所的逻辑卷（β）并写入，为初始复制。

另一是指在初始复制结束后，只把主装置 1 对主场所的逻辑卷（α）进行更新的部分的数据发送给副场所的逻辑卷（β）并写入，为更新复制。

初始复制是把成为复制源的逻辑卷所具有的全部逻辑块的存储内容转发到复制目的地的逻辑卷上的处理，对于执行初始复制的存储控制装置，处理负荷大。因此，在本实施例中，对于防止在初始复制时过大的负荷集中在第一存储控制装置的手段进行说明。

首先在最初，说明在如图所示的前提结构中，主场所的第一存储控制装置 10（SW1）执行逻辑卷的复制的场合。

主副两场所的第一存储控制装置 10（SW1）、10（SW2）分别具备用于取入对应的第二存储控制装置 200（1）、200（2）的逻辑卷的映射表 TM1、TM2。设置在主场所的第一存储控制装置 10（SW1）具备例如图 16（a）所示的映射表 TM1。设置在副场所的第一存储控制装置 10（SW2）具备例如图 16（b）所示的映射表 TM2。在任一映射表中都包含如前所述的向对应的第二存储控制装置 200（1）、200（2）的实 LU222（1）、222（2）的存取端口 ID、卷 ID、地址。

在进行初始复制的场合，主装置 1 生成例如图 16（c）所示的指令（初始复制开始请求）。该指令包含请求开始初始复制的请求 ID、特定复制源的装置的装置 ID、特定复制源的逻辑卷的卷 ID、特定复制目的地的装置的装置 ID、特定复制目的地的逻辑卷的卷 ID。这里，复制源的逻辑卷是主场所的第一存储控制装置 10（SW1）所具有的假想的 LU72（1）的卷（α），复制目的地的逻辑卷是副场所的第一存储控制装置 10（SW2）所具有的假想的 LU72（2）的卷（β）。

参照图 17 说明该处理的流程。图 17 所示的流程图表示通过主场所的第一存储控制装置 10（SW1）的指令控制单元 100（1）执行的处理的概要。主装置 1 向主场所的第一存储控制装置 10（SW1）发送如图 16（c）所示的指令后，指令控制单元 100（1）解析接受的指令，进行数据的读出位置的初始设定（S11）。即，指令控制单元 100（1）根据从主装置 1 接收的指令中的复制源装置 ID 以及复制源卷 ID，特定主场所的第一存储

控制装置 10（SW1）提供的假想的 LU72（1）（卷 ID＝α），把该假想的 LU72（1）的起始逻辑块地址作为数据读出位置进行初始设定。

接下来，指令控制单元 100（1）从初始设定的数据的读出位置读出作为一次的发送数据量而预先决定的量的数据（S12）。这里，指令控制单元 100（1）通过参照映射表 TM，检索分配给假想的 LU72（1）的实 LU222（1），从实 LU222（1）读出一次的量的数据。

然后，指令控制单元 100（1）向从主装置 1 指定的复制目的地、即向设置在副场所的第一存储控制装置 10（SW2）的假想的 LU72（2）（卷 ID＝β）发送刚才读出的数据并使其写入。即，通过通常的写入指令，向副场所的第一存储控制装置 10（SW2）请求数据的写入。通过该写入指令指定的卷 ID 为 β，写入地址为起始逻辑块地址。

从主场所的第一存储控制装置 10（SW1）的启动端口 12（1）发送的写入指令经由通信网络 2（1）、2（2）到达副场所，被副场所的第一存储控制装置 10（SW2）的目标端口 11（2）接收。

第一存储控制装置 10（SW2）的指令控制单元 100（2）解析写入指令，参照映射表 TM2，检索与假想的 LU72（2）（卷 ID ＝β）对应的实 LU222（2）（卷 ID＝β 2）。然后，指令控制单元 100（2），为了把接收的数据写入实 LU222（2），生成数据的写入指令，把该写入指令从启动端口 12（2）发送给第二存储控制装置 200（2）（S13）。第二存储控制装置 200（2）通过目标端口 201（2）接收写入指令后，指令控制单元 210（2）使接收的数据存储在实 LU222（2）的起始逻辑块地址。写入结束后，第二存储控制装置 200（2）向第一存储控制装置 10（SW1）发送写入结束报告。另外，副场所的第一存储控制装置 10（SW2）把写入结束报告发送给主场所的第一存储控制装置 10（SW1）。此外，写入结束报告的发送时刻，根据场合可以采用同步式也可以采用非同步式。

在结束一次的量的写入后，主场所的第一存储控制装置 10（SW1）的指令控制单元 100（1）从上次读出的位置前进一次的数据量的位置，更新数据读出位置（S14）。在超过复制源的逻辑卷的最终逻辑块地址之前，重复以上 S12～S14 的操作，通过该方式，结束双卷之间的初始复制。

这样，在正副两场所之间的双卷的初始复制中，主导的第一存储控制装置 10（SW1）需要多次反复执行 S12～S14 的处理，其负担大。因此，为了降低第一存储控制装置 10 的负荷，提出了进一步改善的方法。

[逻辑卷复制 2]

根据图 18～图 21，说明减轻第一存储控制装置 10 的负荷，执行双卷的初始复制的场合。

和上述同样，正副两场所的第一存储控制装置 10（SW1）、10（SW2）为了把分别对应的第二存储控制装置 200（1）、200（2）所具有的实 LU222（1）、222（2）作为自己的假想的 LU72（1）、72（2）进行利用，分别具备如图 19（a）、（b）所示的映射表 TM1、TM2。该映射表 TM1、TM2 在构成存储系统时登录。另外，如图 19（c）、（d）所示，正副两场所的第一存储控制装置 10（SW1）、10（SW2）分别具备管理对应的第二存储控制装置 200（1）、200（2）支持的功能的全体的功能管理表 TF1、TF2。即，在正副两场所的各场所，第一存储控制装置 10（SW1）、10（SW2）把第二存储控制装置 200（1）、200（2）的实 LU222（1）、222（2）置于间接地支配下的同时，掌握第二存储控制装置 200（1）、200（2）具有的支持功能。

主装置 1，和上述同样，生成例如图 19（e）所示的构造的初始复制开始指令，通过通信网络 2（1）发送给住场所的第一存储控制装置 10（SW1）。在该初始复制开始指令通过目标端口 11（1）被第一存储控制装置 10（SW1）接收后，指令控制单元 100（1）解析接收的指令，使初始复制处理开始。对于以下的动作流程，也在参照图 20 的同时进行说明。图 20 表示通过接收初始复制开始指令的指令控制单元 100（1）执行的处理的概要。

指令控制单元 100（1）最初进行对设置在正副两场所的第二存储控制装置 200（1）、200（2）支持的功能的确认（S21）。指令控制单元 100（1）参照如图 19（c）所示的功能管理表 TF1，确认主场所侧的第二存储控制装置 200（1）是否支持逻辑卷复制功能。例如，将功能 F1 假定为逻辑卷复制功能，由于设定了“可”，因此指令控制单元 100（1）知道可

以使主场所侧的第二存储控制装置 200（1）代行逻辑卷复制处理。

接下来，指令控制单元 100（1）通过通信网络 2（1）、2（2）从启动端口 12（1）向副场所侧的第一存储控制装置 10（SW2）发送请求取得功能管理表 TF2 的指令（表取得请求）。该取得指令通过目标端口 11（2）输入到指令控制单元 100（2）。指令控制单元 100（2）把图 19（d）所示的功能管理表 TF2 的内容作为对应取得指令的响应返回到主场所侧。主场所侧的指令控制单元 100（1）根据从副场所侧取得的功能管理表 TF2 的内容掌握副场所侧的第二存储控制装置 200（2）具备逻辑卷复制功能这一事实。所取得的功能管理表 TF2 的内容存储在第一存储控制装置 10（SW1）的公共存储器或者超高速缓冲存储器中。此外，不限于把功能管理表 TF2 的全部内容从副场所侧发送给主场所侧的场合，例如，也可以从主场所侧的指令控制单元 100（1）对副场所侧的指令控制单元 100（2）问询副场所侧的第二存储控制装置 200（2）是否支持特定的功能。

这样，指令控制单元 100（1）确认设置在主场所和副场所双方的第二存储控制装置 200（1）、200（2）是否分别具备逻辑卷复制功能，在双方都具备逻辑卷复制功能的场合（S22：YES），如下所述的，使通过第二存储控制装置 200（1）、200（2）进行的直接的逻辑卷复制开始。

指令控制单元 100（1）为了从主装置 1 取得与作为复制目的地卷而指定的副场所侧的假想的 LU72(2)对应的实 LU222（2）的信息,对副场所侧的第一存储控制单元 10（SW2）请求取得映射表 TM2（S23）。因为这一映射表取得请求和功能管理表 TF2 的场合同样进行,所以省略其细节。

接下来，指令控制单元 100（1）通过通信网络 2（1）把初始复制开始指令从启动端口 12（1）发送给第二存储控制装置 200（1）的目标端口 201（1）。此时，指令控制单元 100（1）合并从主场所侧的映射表 TM1 抽出来的涉及复制源逻辑卷（α）的信息和从副场所侧的映射表 TM2 抽出来的涉及复制目的地逻辑卷（β）的信息，发送给主场所侧的第二存储控制装置 200（1）。在通过第二存储控制装置 200（1）进行初始复制期间，通过主装置 1 更新假想的 LU72（1）的内容的场合，将涉及该更

新的逻辑块的信息存储在差分位图表 TB 中。

在主场所侧的第二存储控制装置 200（1）通过目标端口 201（1）接收初始复制开始指令后，该指令被指令控制单元 210（1）解析，开始如图 21 所示的初始复制处理。

图 21 表示通过主场所侧的第二存储控制装置 200（1）的指令控制单元 210（1）执行的初始复制处理。指令控制单元 210（1）根据通过初始复制开始指令指定的复制源卷 ID 和在 S24 取得的各逻辑卷的信息，将为复制源的假想的 LU72（1）（卷 ID=α）的起始逻辑块地址初始设定为数据读出位置（S31）。

接下来，指令控制单元 210（1）从设定的数据的读出位置读出预先作为一次的发送数据量设定的量的数据（S32）。此时，指令控制单元 210（1）根据先前从第一存储控制单元 10（SW1）取得的映射表 TM1 的信息检索与假想的 LU72（1）对应的实 LU222（1），从实 LU222（1）（卷 ID=α2）的起始逻辑块地址读出一次发送量的数据。

然后，指令控制单元 210（1）对通过初始复制开始指令指定的复制目的地发送并写入读出的数据（S33）。该写入请求通过通常的写入指令进行。此时，指令控制单元 210（1）根据从第一存储控制单元 10（SW1）取得的映射表 TM2 的信息，向与副场所侧的假想的 LU72（2）对应的实 LU222（2）（卷 ID=β2）设定写入指令中的写入目的地。写入目的地的地址和从实 LU222（1）读出数据时指定的地址相同。

从第二存储控制装置 200（1）的启动端口 202（1）发送的写入指令，经由通信网络 2（1）、2（2），被副场所侧的第二存储控制装置 200（2）的目标端口 201（2）接收。第二存储控制装置 200（2）的指令控制单元 210（2），遵从接受的写入指令将接收的数据写入实 LU222（2）的规定的位置，并向主场所侧的第二存储控制装置 200（1）报告写入结束。

在结束一次的写入指令的处理后，指令控制单元 210（1）通过从上次的读出位置前进相当于一次发送数据量的位置，更新数据读出位置（S34）。指令控制单元 210（1），在把实 LU222（1）的存储内容全部移动到实 LU222（2）之前，重复执行 S32～S34 的处理。在初始复制结束

后，指令控制单元 210（1）向第一存储控制单元 10（SW1）报告初始复制结束。

在初始复制结束的场合返回到图 20，第一存储控制单元 10（SW1）的指令控制单元 100（1）参照差分位图表 TB。在差分位图表 TB 中，保持有涉及在初始复制中主装置 1 对为本来的复制源的假想的 LU72（1）进行的新数据的写入的信息，即被更新的逻辑块的地址信息。

在通过参照差分位图表 TB、检测出在初始复制中假想的 LU72（1）的存储内容被变更的场合，指令控制单元 100（1）执行使在差分位图表 TB 中表示的全部的逻辑块写入复制目的地的假想的 LU72（2）的处理（S25）。该被更新的数据的复制，例如可以按照[逻辑卷复制 1]中叙述的方法进行。

这样，进行将在初始复制中写入的新数据写入副场所侧的处理，是根据下面的理由。通过第二存储控制装置 200（1）中的指令控制单元 210（1）进行的初始复制，从逻辑块地址小的到大的依次进行。例如，在初始复制接近几乎结束的时候，在逻辑卷的起始逻辑块地址附近进行新数据的写入的话，指令控制单元 210（1）不能将该数据复制在副场所侧。这样，为了避免可能发生的未复制的逻辑块，进行根据差分位图表 TB 的新写入数据的复制。

此外，在图 20 中，在判断为正副两场所的第二存储控制装置 200（1）、200（2）的任何一方不支持逻辑卷复制功能的场合（S22：NO），如在[逻辑卷复制 1]中所述的，主场所侧的第一存储控制装置 10（SW1）主导进行初始复制（S26）。

根据这样构成的本实施例，和上述实施例同样，可以把从主装置 1 请求的数据处理（主场所和副场所之间的逻辑卷复制）委托给第二存储控制装置 200（1）、200（2），可以使数据处理的负荷分散，减轻第一存储控制装置 10（SW1）、10（SW2）的负荷。因此，可以把第一存储控制装置 10（SW1）的信息处理能力用于实现其它的服务，可以使存储系统高效运行。

【实施例 3】

接下来，参照图 22，说明本发明的第三实施例。本实施例的特征在于，第一存储控制装置 10 具备多个假想的 LU72（1）、72（2），适用于使这些多个假想的 LU72（1）、72（2）的存储内容同步化的场合。

第一存储控制装置 10 例如由智能化的纤维通道开关构成，具有两个假想的 LU72（1）、72（2）。一方的假想的 LU72（1）是主卷、另一方的假想的 LU72（2）是辅助卷。各假想的 LU72（1）、72（2）的实体是第二存储控制装置 200（1）、200（2）的实 LU222（1）、222（2）。

在将假想的 LU72（1）的存储内容复制到假想的 LU72（2）的场合，可以与在第二实施例的[逻辑卷复制 1]或者[逻辑卷复制 2]中叙述的方法同样进行。

例如，根据一种方法，第一存储控制装置 10 通过从起始逻辑块地址每次按规定量读出第二存储控制装置 200 的实 LU222（1）的数据，把读出的数据写入第二存储控制装置 200 的实 LU222（2）的规定位置，可以把实 LU222（1）的全部数据复制到实 LU222（2）。

另外，根据另外的方法，判别第二存储控制装置 200 是否支持镜像功能，在具有镜像功能的场合，从第一存储控制装置 10 向第二存储控制装置 200 发送镜像开始指令。在该镜像开始指令中，至少包含复制源卷 ID（α）和复制目的地卷 ID（β）。此时，可以一起发送映射表 TM 的存储内容。或者，也可以通过参照映射表 TM，分别将复制源卷 ID 变换为 α 2、将复制目的地卷 ID 变换为 β 2，并发送镜像开始指令。

在第二存储控制装置 200 接收镜像开始指令后，将从实 LU222（1）的起始逻辑块地址到最终逻辑块地址的全部数据每次以规定量复制到实 LU222（2）。此外，在镜像中，在数据被主装置 1 更新的场合，和上面说明的同样，可以利用差分位图表，从后边复制被更新了的数据。

此外，本发明不限定于上述实施例。本技术领域的人可以在本发明的范围内进行各种追加和变更等。在上述各实施例中，中心说明了智能化的纤维通道开关，但是本发明不限于此，也可以广泛适用于磁盘阵列装置等。另外，也可以适用于在第一存储控制装置内设置与各种不同的第二存储控制装置的实存储区域对应的假想的存储区域的场合。

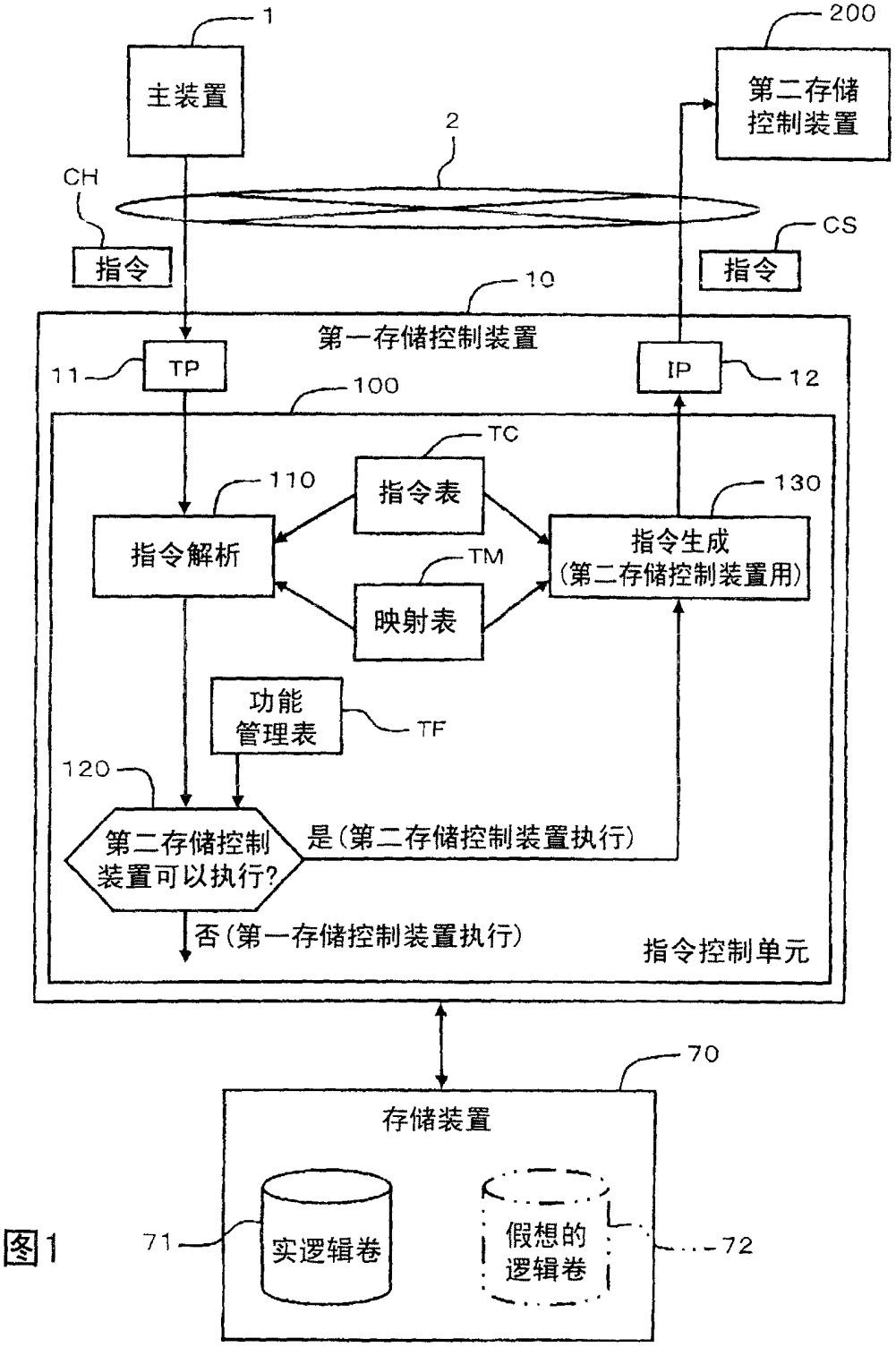

1
主装置
200
第二存储
控制装置
2
CH
指令
CS
指令
10
第一存储控制装置
11
TP
IP
12
100
TC
指令表
110
130
指令解析
指令生成
(第二存储控制装置用)
TM
映射表
功能
管理表
TF
120
第二存储控制
装置可以执行?
是(第二存储控制装置执行)
否(第一存储控制装置执行)
指令控制单元
70
存储装置
71
实逻辑卷
假想的
逻辑卷
72

图1

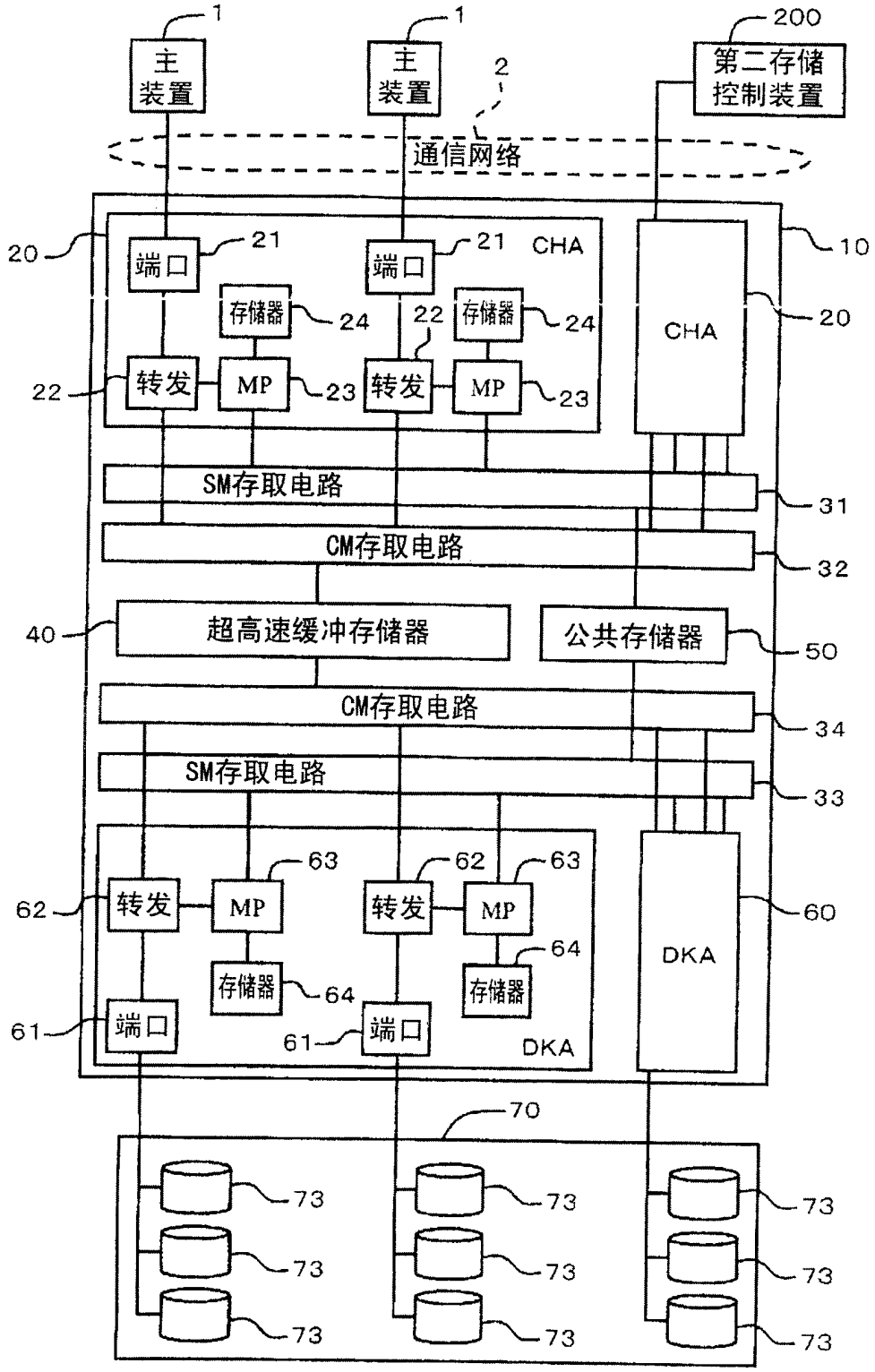

1
主装置
1
主装置
2
200
第二存储控制装置
通信网络
20
端口
21
端口
21
CHA
10
存储器
24
22
存储器
24
CHA
20
22
转发
MP
23
转发
MP
23
SM存取电路
31
CM存取电路
32
40
超高速缓冲存储器
公共存储器
50
CM存取电路
34
SM存取电路
33
63
62
63
62
转发
MP
转发
MP
60
64
DKA
存储器
64
存储器
61
端口
61
端口
DKA
70
73
73
73
73
73
73
73
73
73

图2

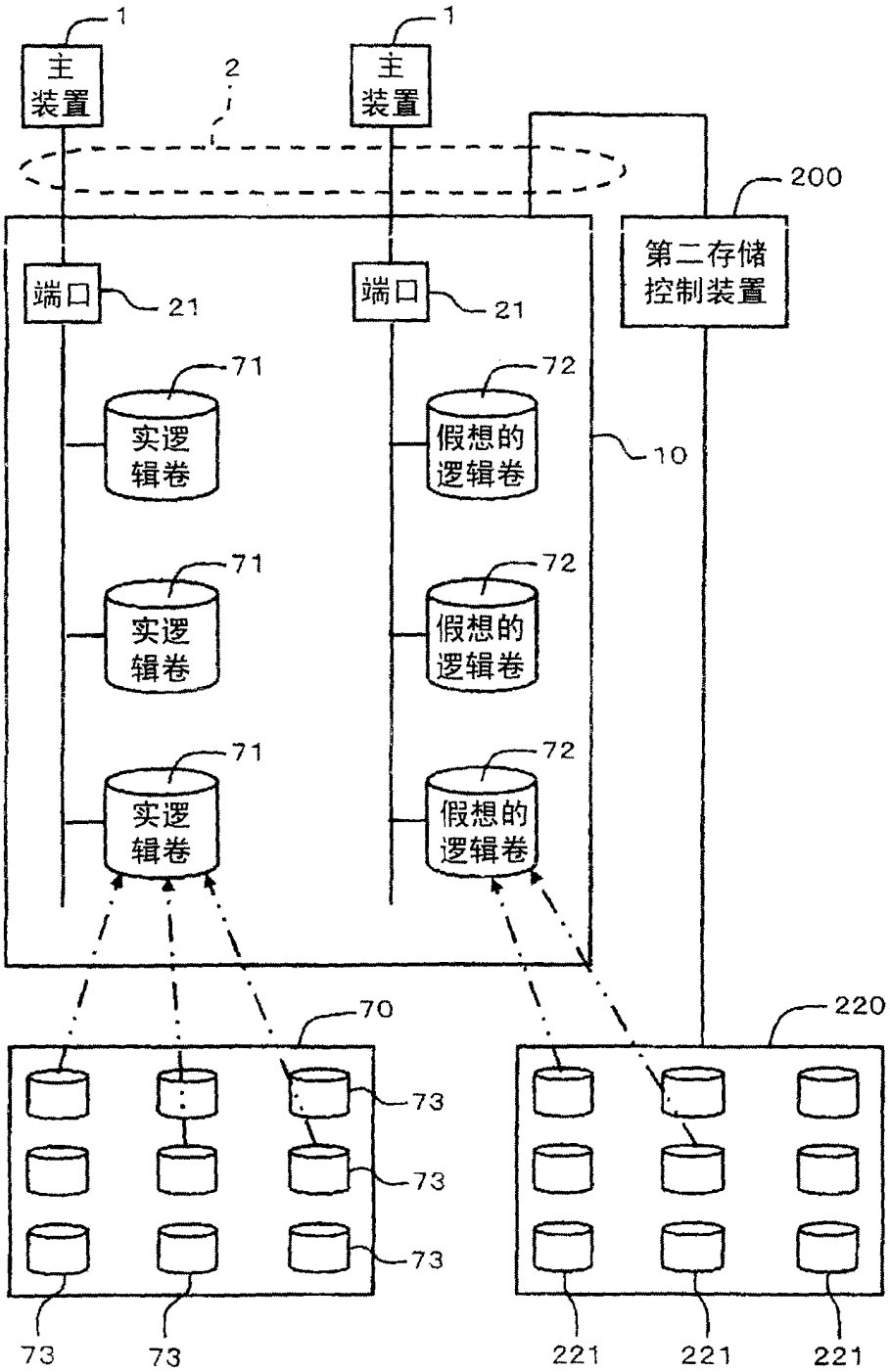

1
主装置
2
1
主装置
200
第二存储控制装置
端口
21
端口
21
71
实逻辑卷
72
假想的逻辑卷
10
71
实逻辑卷
72
假想的逻辑卷
71
实逻辑卷
72
假想的逻辑卷
70
220
73
73
73
73
73
221
221
221

图3

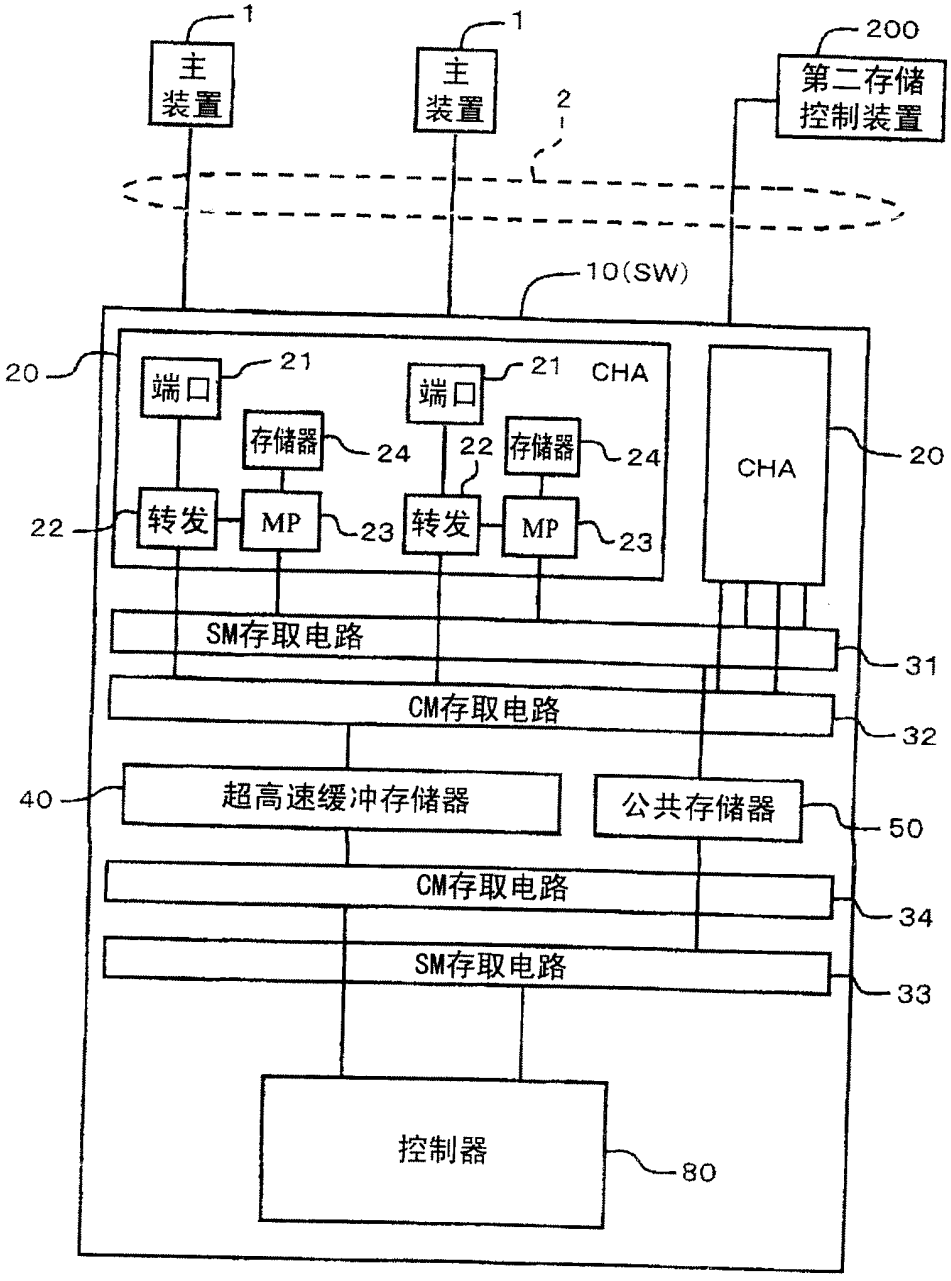

1
主装置
1
主装置
2
200
第二存储控制装置
10(SW)
20
端口
21
端口
21
CHA
存储器
24
存储器
24
22
转发
MP
23
转发
MP
23
22
CHA
20
SM存取电路
31
CM存取电路
32
40
超高速缓冲存储器
公共存储器
50
CM存取电路
34
SM存取电路
33
控制器
80

图4

图5

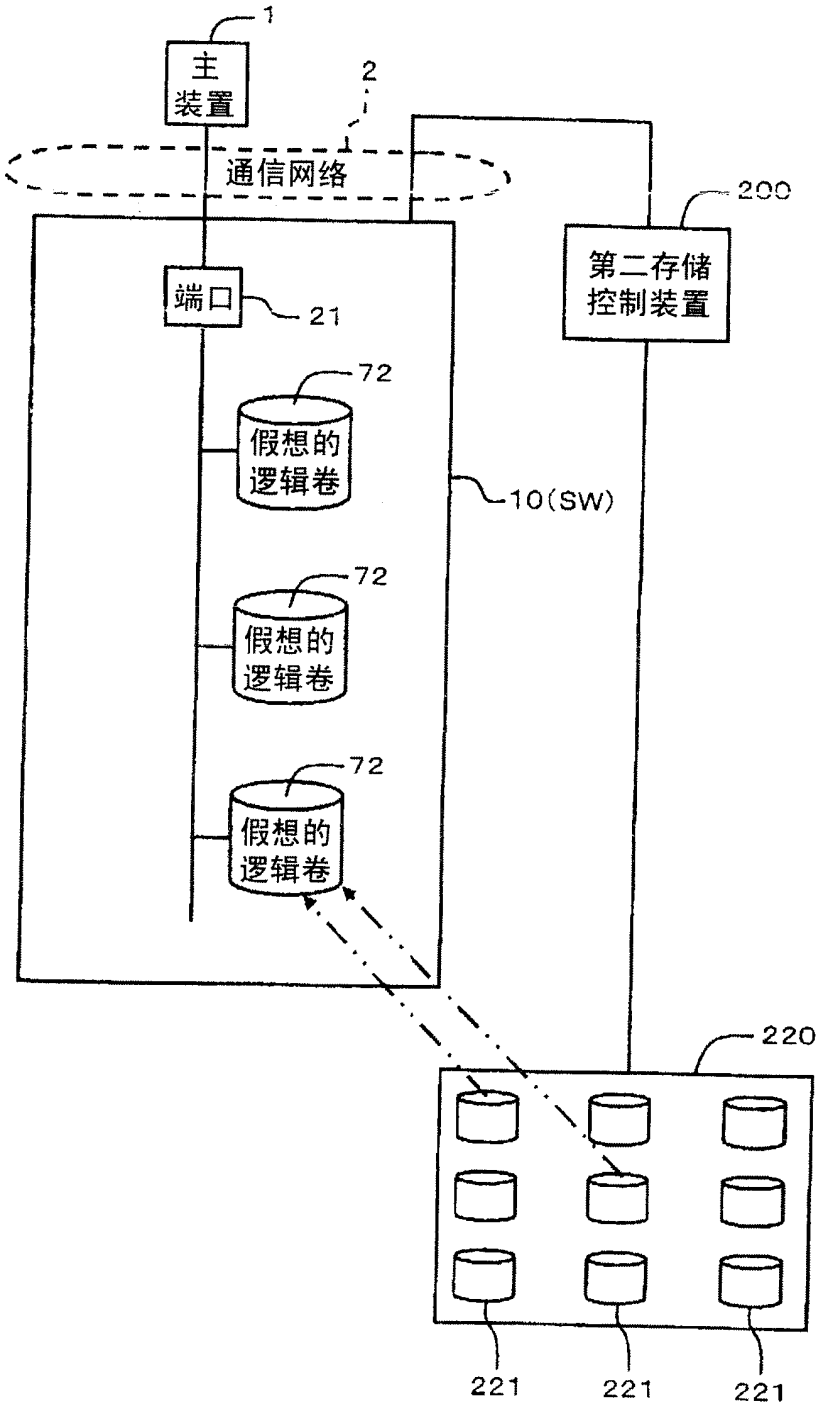
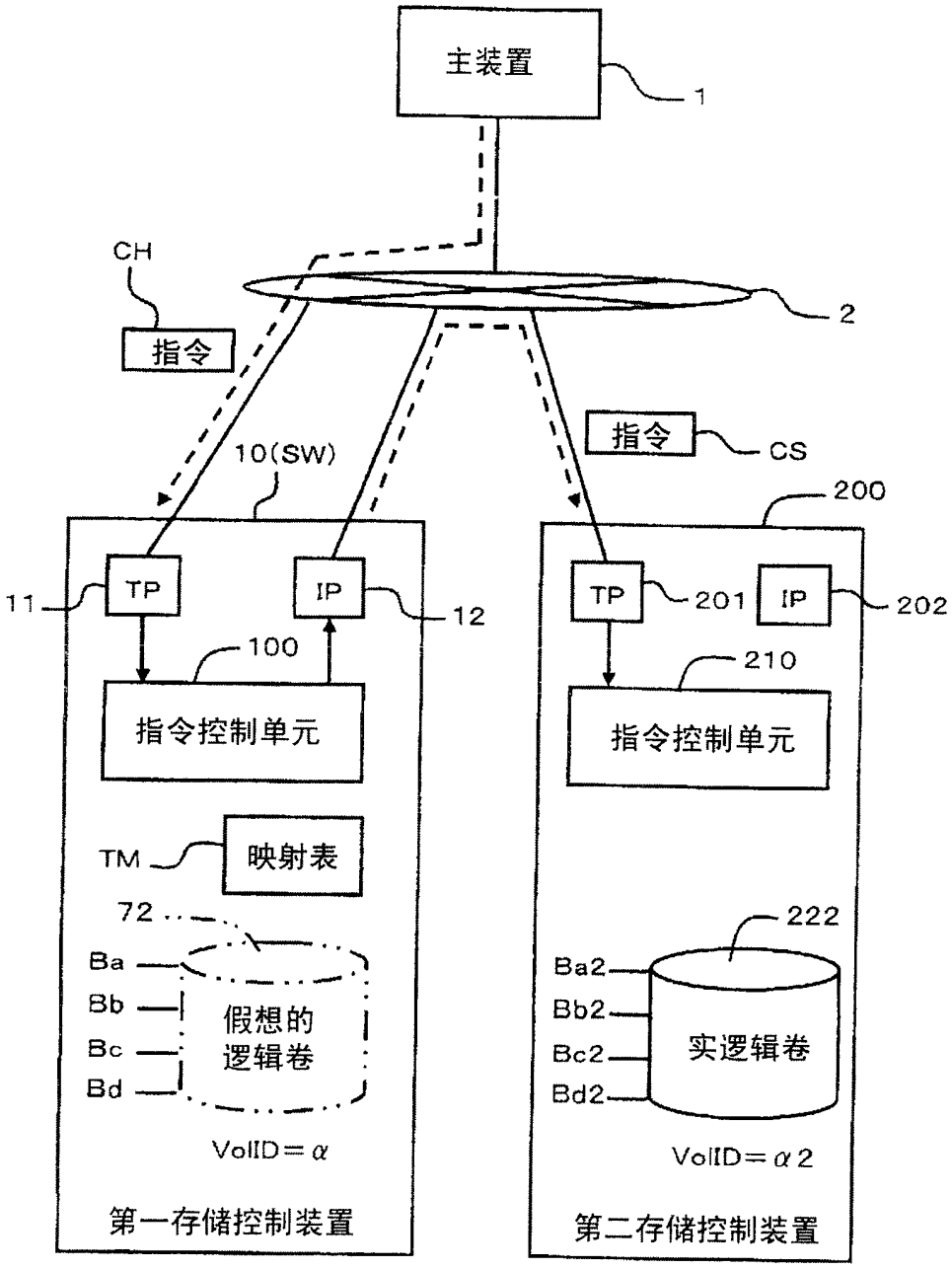

1
主
装置
2
通信网络
200
第二存储
控制装置
端口
21
72
假想的
逻辑卷
10(SW)
72
假想的
逻辑卷
72
假想的
逻辑卷
220
221
221
221

图6


主装置
1
CH
指令
2
10(SW)
指令
CS
200
11
TP
IP
12
TP
201
IP
202
100
指令控制单元
210
指令控制单元
TM
映射表
72
Ba
Bb
Bc
Bd
假想的
逻辑卷
VolID＝α
222
Ba2
Bb2
Bc2
Bd2
实逻辑卷
VolID＝α2
第一存储控制装置
第二存储控制装置

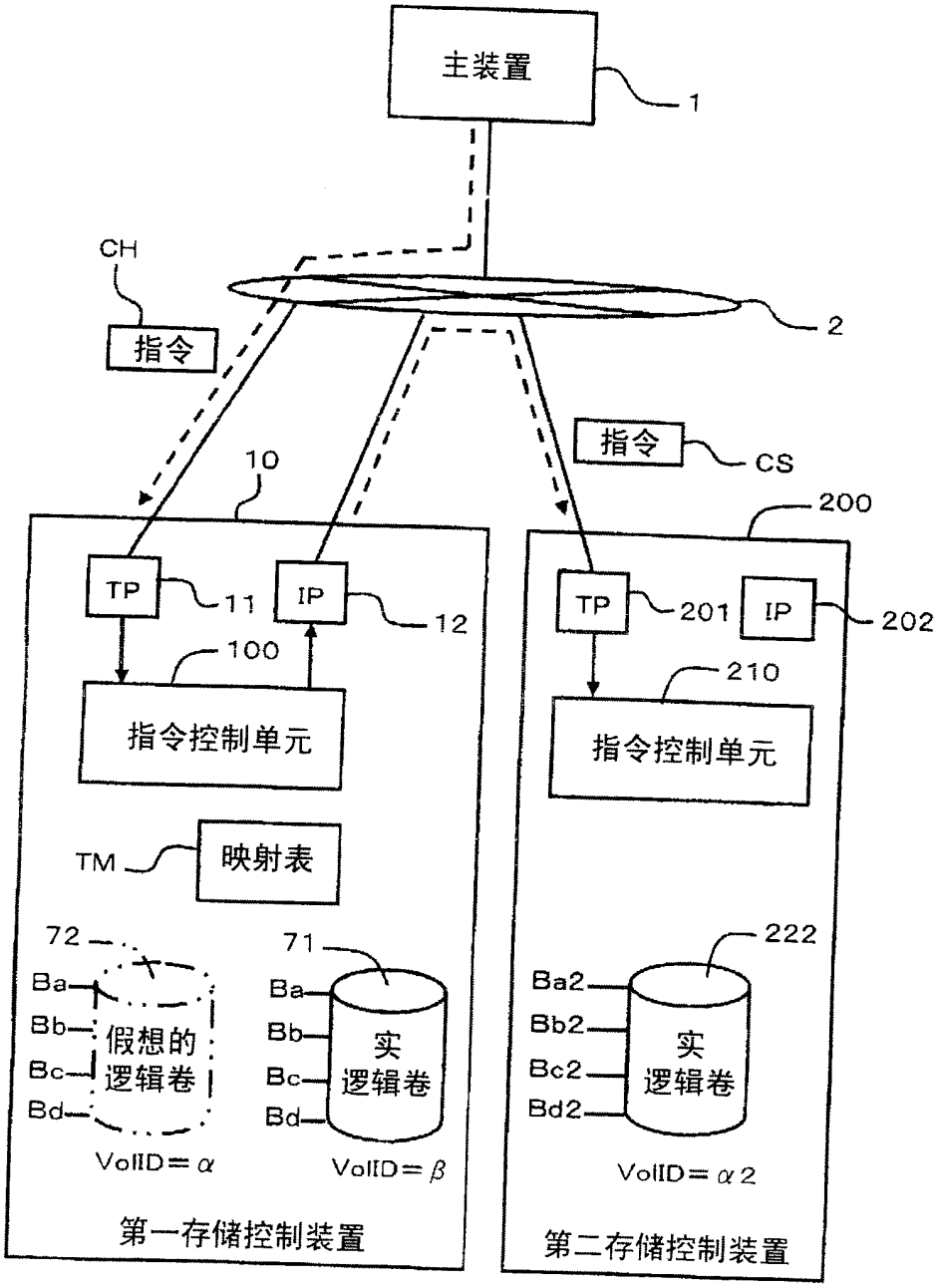

主装置
1
CH
指令
2
指令
CS
10
200
TP
11
IP
12
TP
201
IP
202
100
210
指令控制单元
指令控制单元
TM
映射表
72
71
222
Ba
Bb
Bc
Bd
假想的
逻辑卷
实
逻辑卷
Ba2
Bb2
Bc2
Bd2
实
逻辑卷
VolID＝α
VolID＝β
VolID＝α2
第一存储控制装置
第二存储控制装置

图7

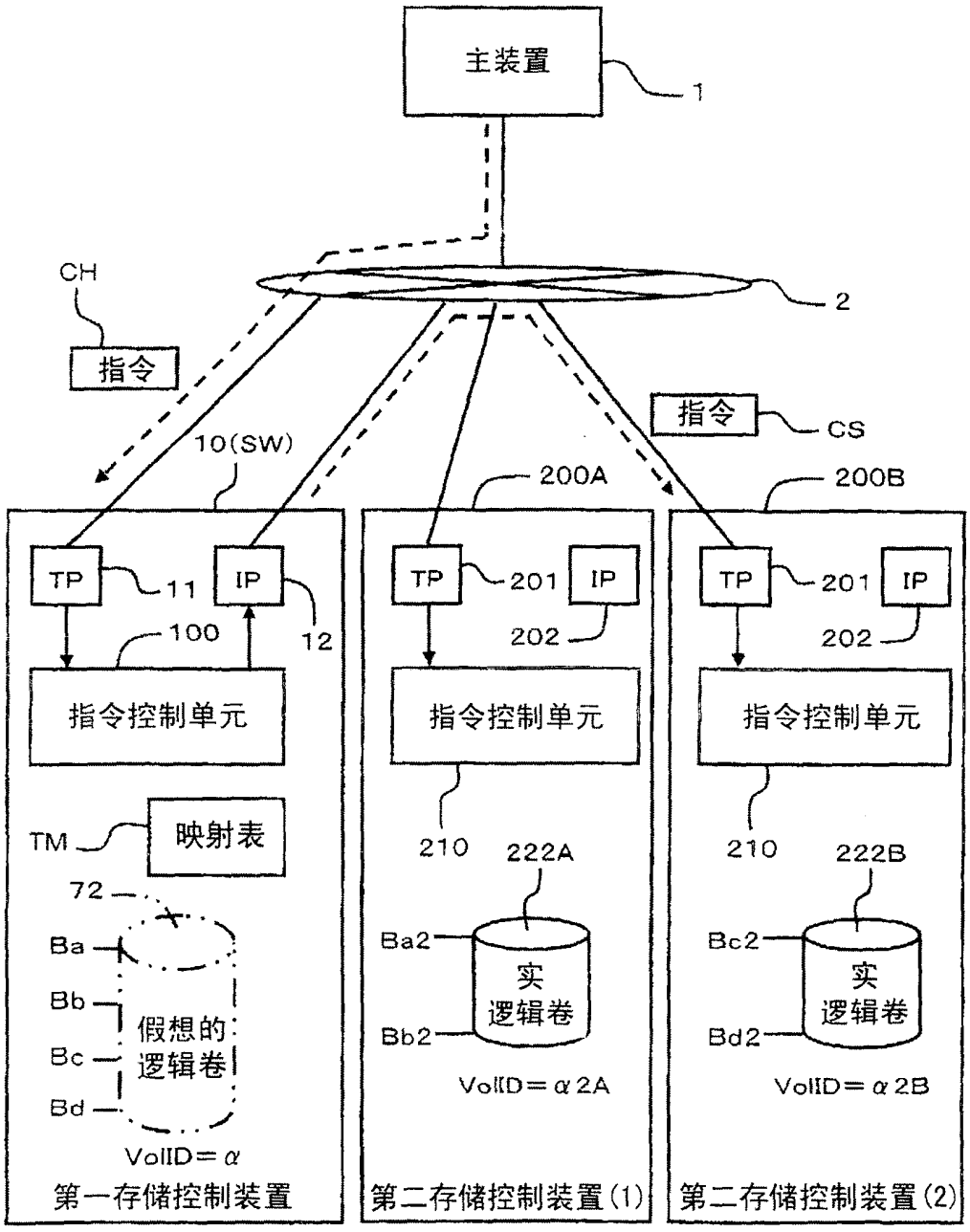

主装置
1
2
CH
指令
10(SW)
指令
CS
200A
200B
TP
11
IP
12
100
TP
201
IP
202
TP
201
IP
202
指令控制单元
指令控制单元
指令控制单元
TM
映射表
72
Ba
Bb
Bc
Bd
假想的
逻辑卷
VolID=α
210
222A
Ba2
Bb2
实
逻辑卷
VolID=α2A
210
222B
Bc2
Bd2
实
逻辑卷
VolID=α2B
第一存储控制装置
第二存储控制装置(1)
第二存储控制装置(2)

图8

(a) 映射表的例子(纤维通道开关的场合)

Vol ID	BLK ADDR	装置ID	Port ID	Vol ID	BLK ADDR
α	Ba	SD2	TP2	α2	Ba2
⋮	⋮	⋮	⋮	⋮	⋮
α	Bd	SD2	TP2	α2	Bd2

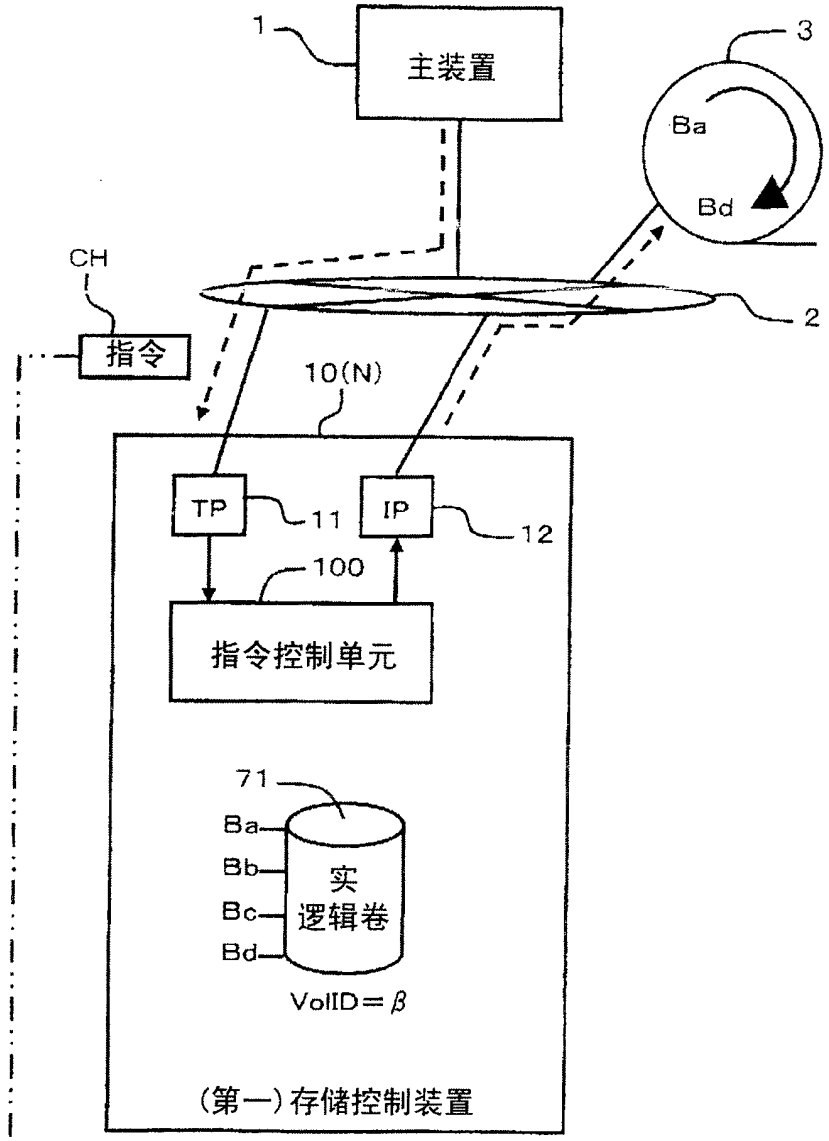
(b) 映射表的例子(磁盘阵列装置的场合)

Vol ID	BLK ADDR	装置ID	Port ID	Vol ID	BLK ADDR
α	Ba	SD2	TP2	α2	Ba2
⋮	⋮	⋮	⋮	⋮	⋮
α	Bd	SD2	TP2	α2	Bd2
β	Ba	SD1	INTERNAL	—	—
⋮	⋮	⋮	⋮	⋮	⋮
β	Bd	SD1	INTERNAL	—	—

(c) 映射表的例子(第二存储控制装置有多个场合)

Vol ID	BLK ADDR	装置ID	Port ID	Vol ID	BLK ADDR
α	Ba	SD2(1)	TP2(1)	α2A	Ba2
⋮	Bb	SD2(1)	TP2(1)	α2A	Bb2
⋮	Bc	SD2(2)	TP2(2)	α2B	Bc2
α	Bd	SD2(2)	TP2(2)	α2B	Bd2

图9



(a) 备份数据的场合

指令代码 (复制请求)	复制源地址 (Ba－Bd)	复制目的地装置ID (备份装置)

(b) 在备份装置中写入控制信息的场合

指令代码 (写入请求)	控制信息	复制目的地装置ID (备份装置)

图10

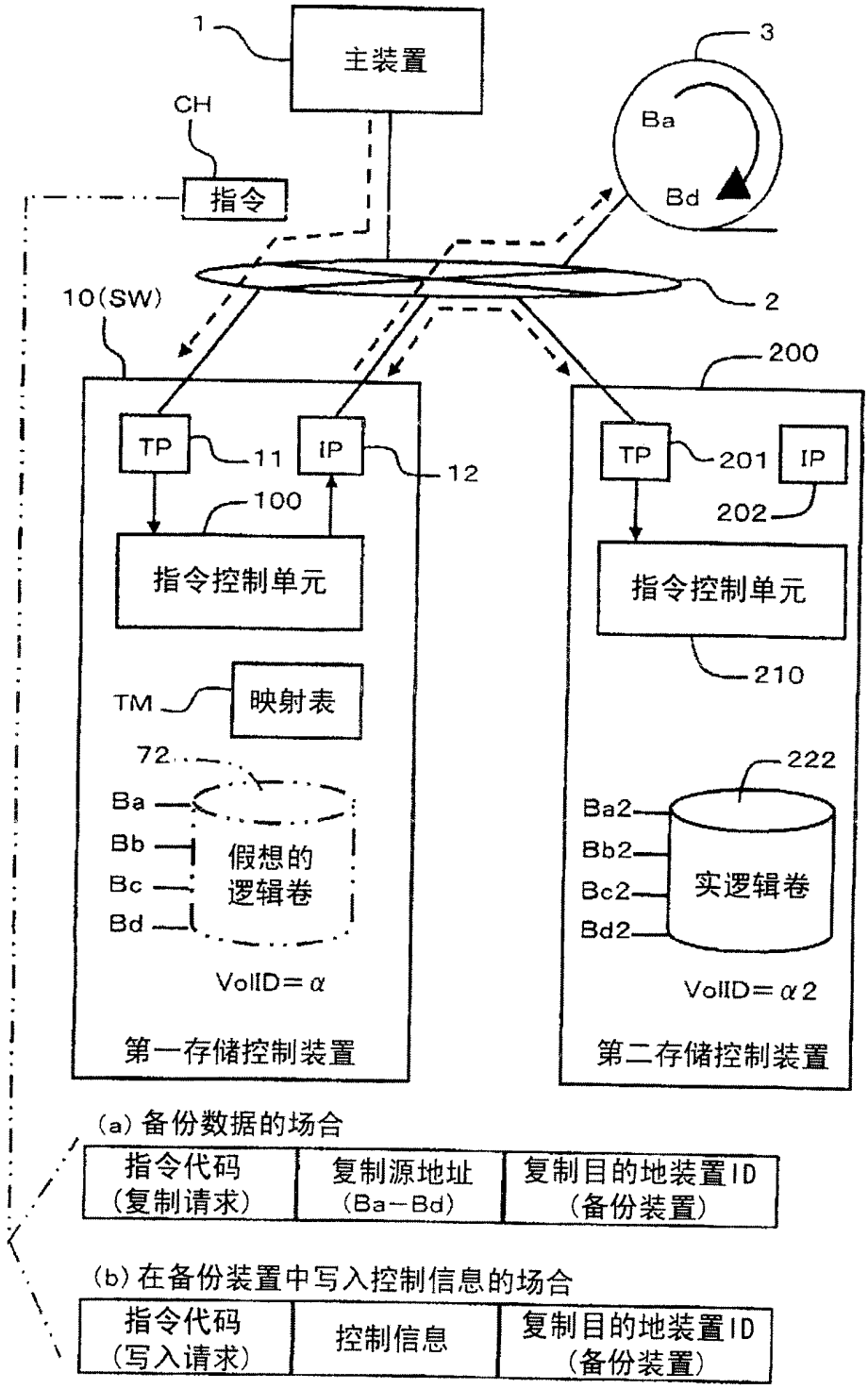

1
主装置
CH
指令
3
Ba
Bd
2
10(SW)
200
TP
11
IP
12
100
指令控制单元
TM
映射表
72
Ba
Bb
Bc
Bd
假想的
逻辑卷
VolID=α
第一存储控制装置
TP
201
IP
202
指令控制单元
210
222
Ba2
Bb2
Bc2
Bd2
实逻辑卷
VolID=α2
第二存储控制装置
(a) 备份数据的场合
指令代码
(复制请求)
复制源地址
(Ba-Bd)
复制目的地装置ID
(备份装置)
(b) 在备份装置中写入控制信息的场合
指令代码
(写入请求)
控制信息
复制目的地装置ID
(备份装置)

图11

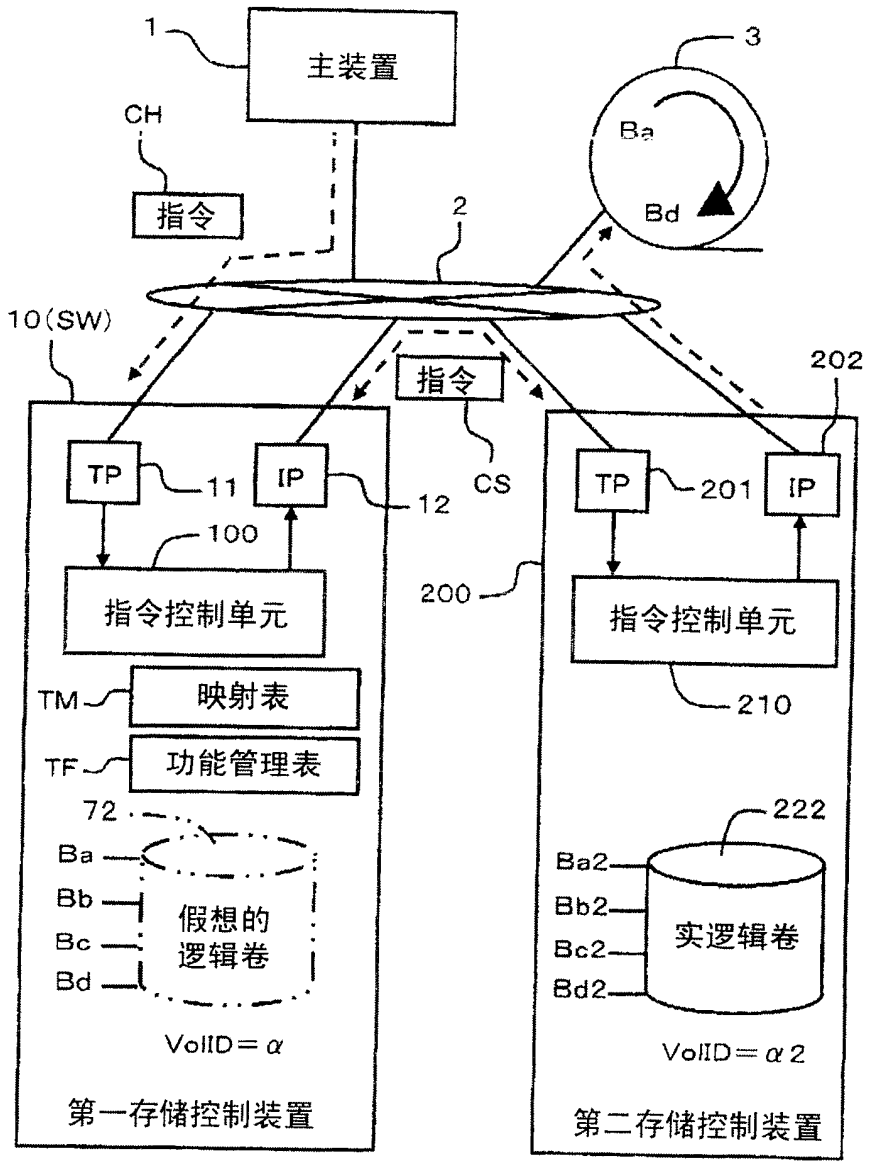

1
主装置
CH
指令
3
Ba
Bd
2
10(SW)
指令
CS
202
TP
11
IP
12
TP
201
IP
100
指令控制单元
200
指令控制单元
210
TM
映射表
TF
功能管理表
72
Ba
Bb
Bc
Bd
假想的
逻辑卷
VolID＝α
第一存储控制装置
222
Ba2
Bb2
Bc2
Bd2
实逻辑卷
VolID＝α2
第二存储控制装置

图12

图13

(a) 功能管理表的例子

TF

装置ID	端口ID	支持功能 F1	支持功能 F2	…	支持功能 Fn
SD2(1)	TP2(1)	可	不可	…	可
⋮	⋮	⋮	⋮	⋮	⋮
SD2(n)	TP2(n)	可	可	…	不可

(b) 主装置→第一存储控制装置

CH

指令代码 (复制请求)	复制源地址 (Ba-Bd)	复制目的地装置ID (备份装置)

(c) 第一存储控制装置→第二存储装置

CS

指令代码 (复制请求)	复制源地址 (Ba2-Bd2)	复制目的地装置ID (备份装置)

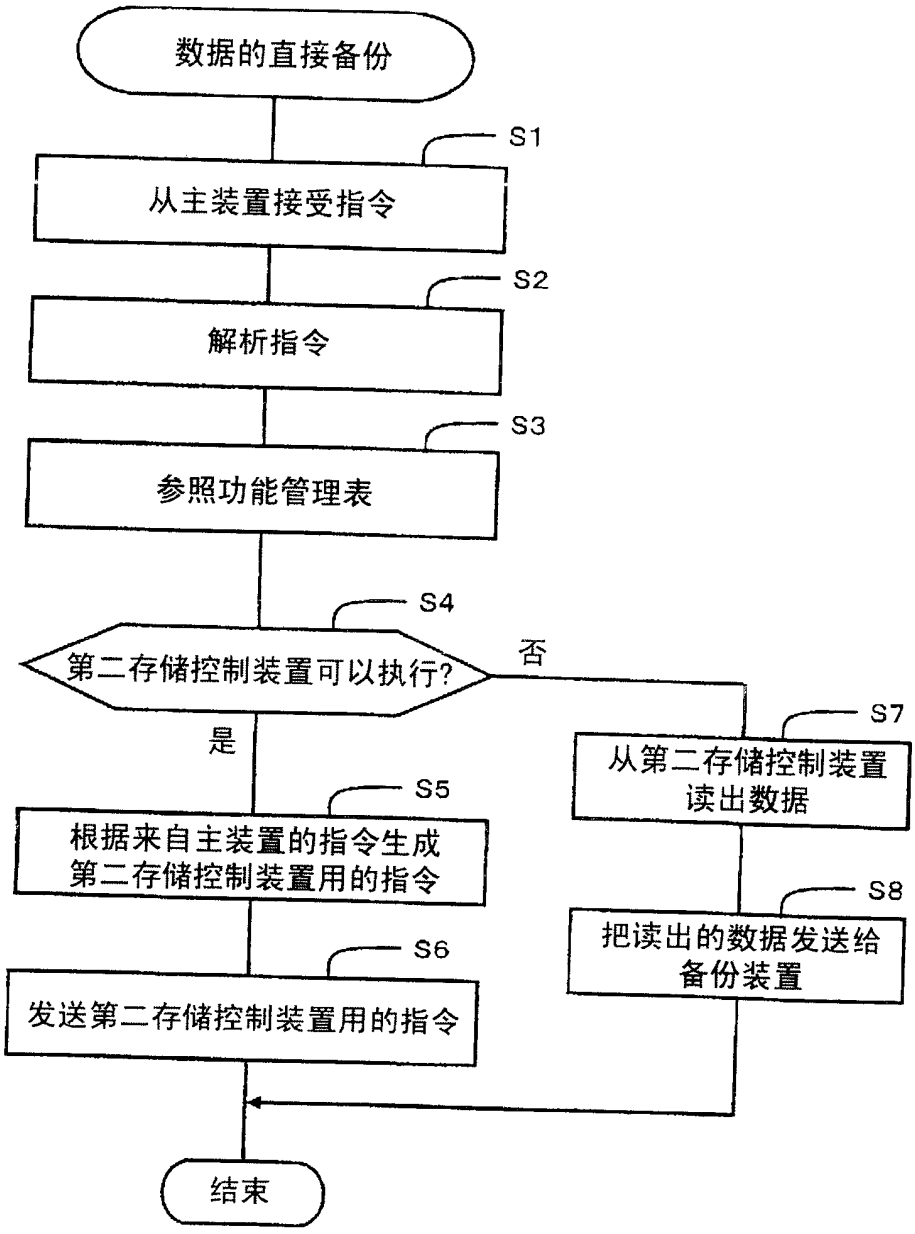

数据的直接备份
S1
从主装置接受指令
S2
解析指令
S3
参照功能管理表
S4
第二存储控制装置可以执行?
否
是
S5
根据来自主装置的指令生成
第二存储控制装置用的指令
S6
发送第二存储控制装置用的指令
S7
从第二存储控制装置
读出数据
S8
把读出的数据发送给
备份装置
结束

图14

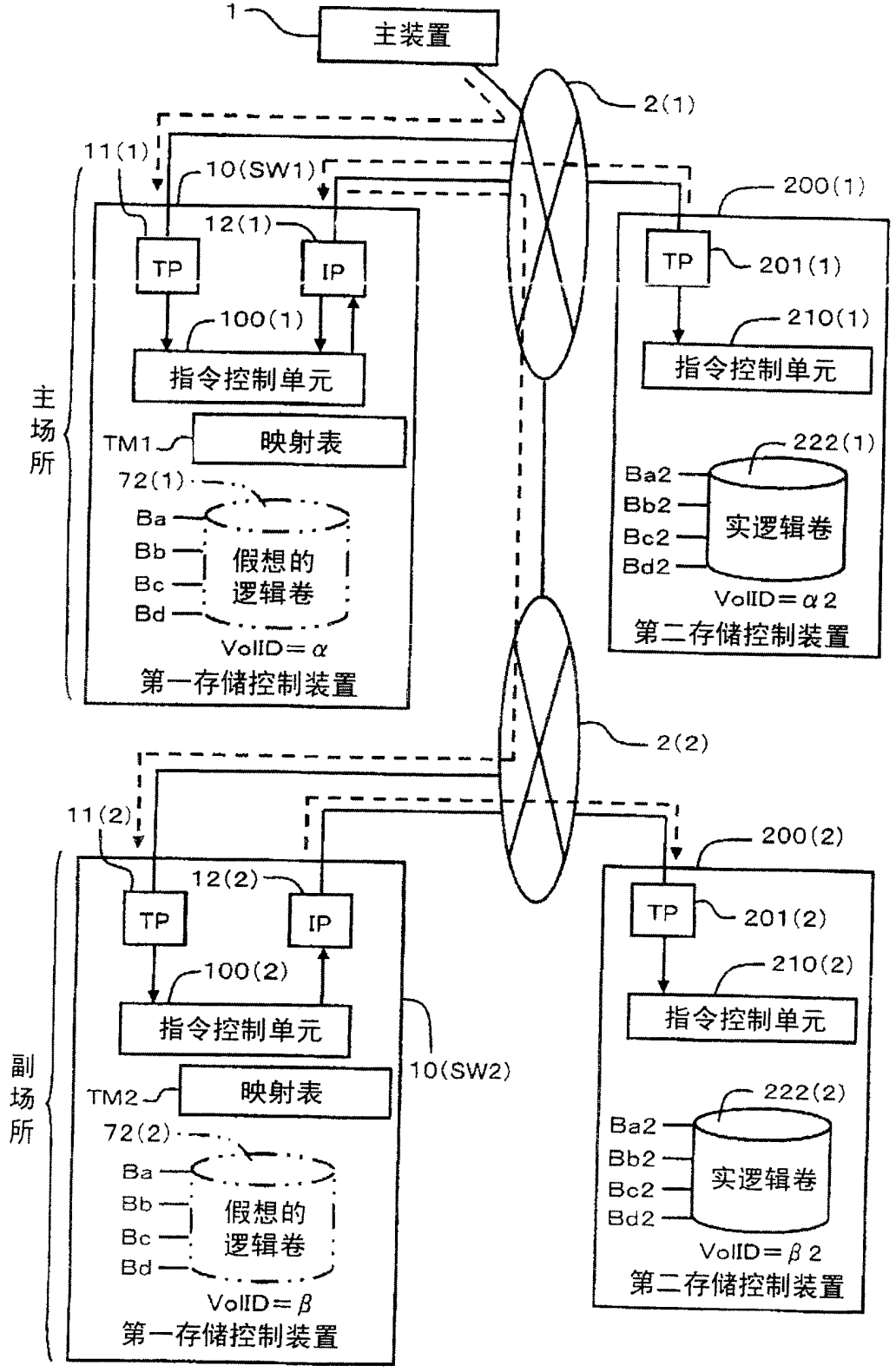

1
主装置
2(1)
11(1)
10(SW1)
200(1)
12(1)
TP
IP
TP
201(1)
100(1)
210(1)
指令控制单元
指令控制单元
主场所
TM1
映射表
222(1)
72(1)
Ba
Bb
Bc
Bd
假想的逻辑卷
Ba2
Bb2
Bc2
Bd2
实逻辑卷
VolID=α
VolID=α2
第一存储控制装置
第二存储控制装置
2(2)
11(2)
200(2)
12(2)
TP
IP
TP
201(2)
100(2)
210(2)
指令控制单元
指令控制单元
10(SW2)
副场所
TM2
映射表
222(2)
72(2)
Ba
Bb
Bc
Bd
假想的逻辑卷
Ba2
Bb2
Bc2
Bd2
实逻辑卷
VolID=β
VolID=β2
第一存储控制装置
第二存储控制装置

图15

(a) 主场所侧映射表的例子

Vol ID	BLK ADDR	装置 ID	端口 ID	Vol ID	BLK ADDR
α	Ba	SD2(1)	TP2(1)	α2	Ba2
⋮	⋮	⋮	⋮	⋮	⋮
α	Bd	SD2(1)	TP2(1)	α2	Bd2

(b) 副场所侧映射表的例子

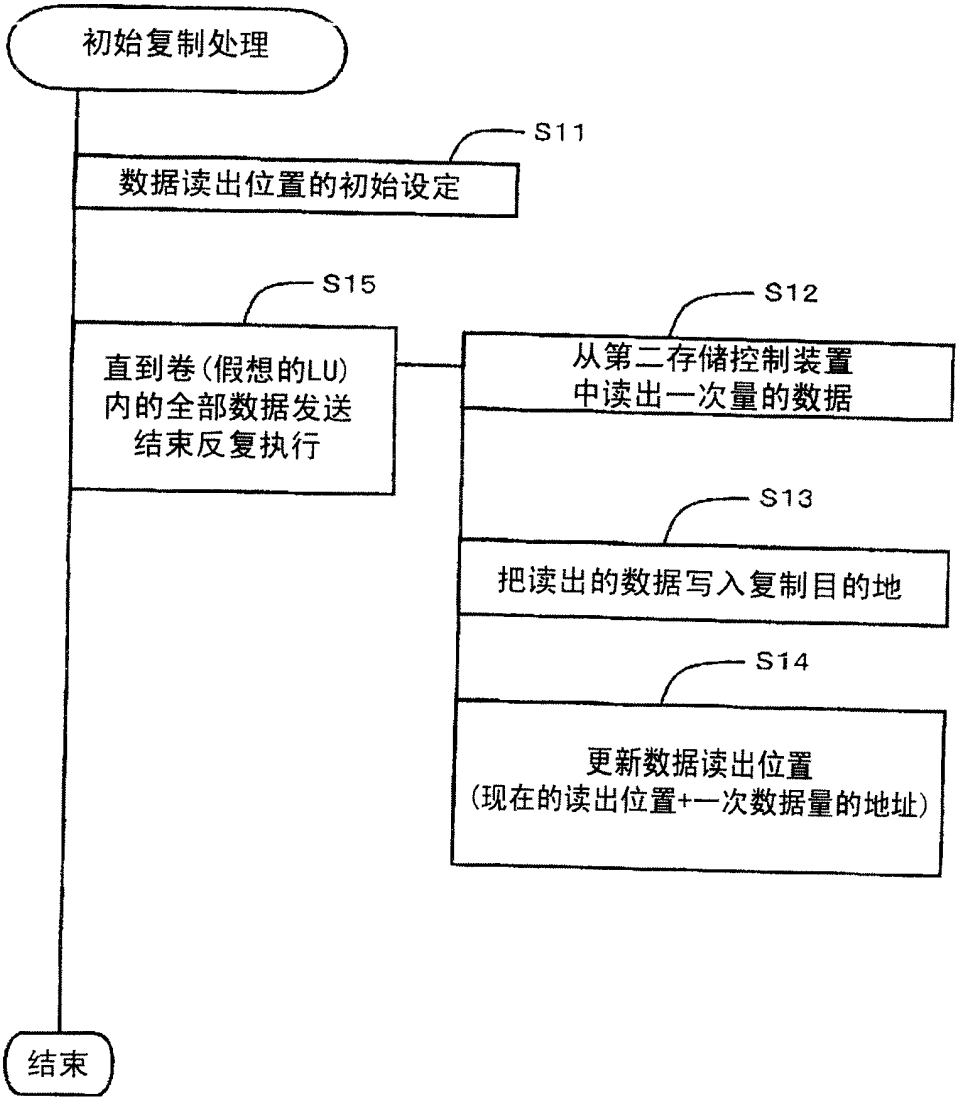
Vol ID	BLK ADDR	端口 ID	端口 ID	Vol ID	BLK ADDR
β	Ba	SD2(2)	TP2(2)	β2	Ba2
⋮	⋮	⋮	⋮	⋮	⋮
β	Bd	SD2(2)	TP2(2)	β2	Bd2

(c) 执行初始复制场合的指令结构

指令代码 (初始复制开始请求)	复制源装置ID (SD1(1))	复制源卷ID (α)	复制目的地装置ID (SD1(2))	复制目的卷ID (β)

图16

图17


初始复制处理
S11
数据读出位置的初始设定
S15
直到卷(假想的LU)
内的全部数据发送
结束反复执行
S12
从第二存储控制装置
中读出一次量的数据
S13
把读出的数据写入复制目的地
S14
更新数据读出位置
(现在的读出位置+一次数据量的地址)
结束

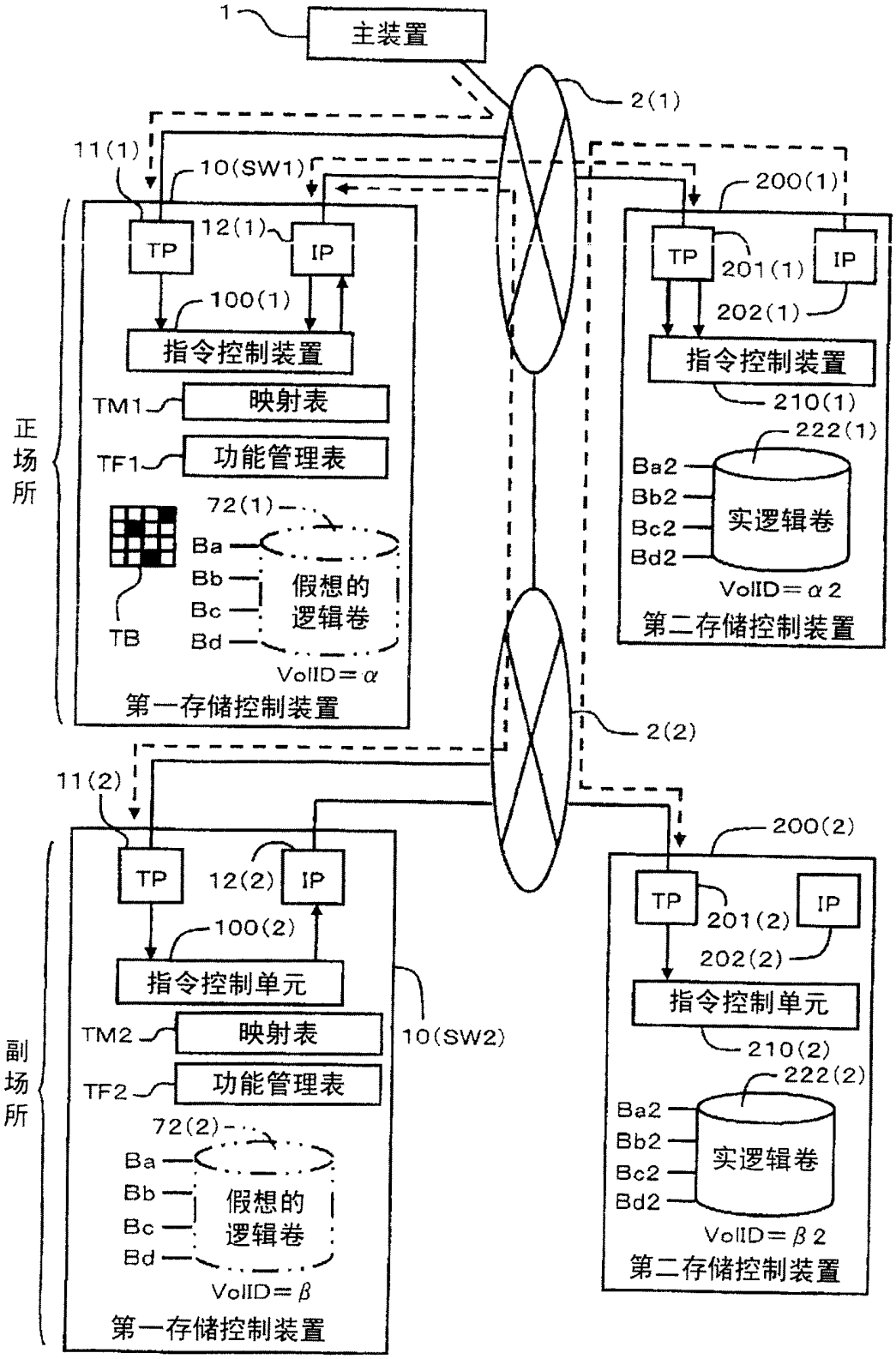

1
主装置
2(1)
11(1)
10(SW1)
200(1)
TP
12(1)
IP
TP
201(1)
IP
100(1)
202(1)
指令控制装置
指令控制装置
210(1)
TM1
映射表
222(1)
正场所
TF1
功能管理表
Ba2
Bb2
Bc2
Bd2
实逻辑卷
72(1)
Ba
Bb
Bc
Bd
假想的逻辑卷
TB
VolID=α2
VolID=α
第二存储控制装置
第一存储控制装置
2(2)
11(2)
200(2)
TP
12(2)
IP
TP
201(2)
IP
100(2)
202(2)
指令控制单元
指令控制单元
210(2)
TM2
映射表
10(SW2)
222(2)
副场所
TF2
功能管理表
Ba2
Bb2
Bc2
Bd2
实逻辑卷
72(2)
Ba
Bb
Bc
Bd
假想的逻辑卷
VolID=β2
VolID=β
第二存储控制装置
第一存储控制装置

图18

(a) 主场所侧映射表的例子

Vol ID	BLK ADDR	装置 ID	端口 ID	Vol ID	BLK ADDR
α	Ba	SD2(1)	TP2(1)	α2	Ba2
⋮	⋮	⋮	⋮	⋮	⋮
α	Bd	SD2(1)	TP2(1)	α2	Bd2

(b) 副场所侧映射表的例子

Vol ID	BLK ADDR	装置 ID	端口 ID	Vol ID	BLK ADDR
β	Ba	SD2(2)	TP2(2)	β2	Ba2
⋮	⋮	⋮	⋮	⋮	⋮
β	Bd	SD2(2)	TP2(2)	β2	Bd2

(c) 主场所侧功能管理表的例子

装置 ID	端口 ID	支持功能 F1	支持功能 F2	...	支持功能 Fn
SD2(1)	TP2(1)	可	不可	...	可

(d) 副场所侧功能管理表的例子

装置 ID	端口 ID	支持功能 F1	支持功能 F2	...	支持功能 Fn
SD2(2)	TP2(2)	可	可	...	不可

(e) 执行初始复制场合的指令结构

指令代码 (初始复制开始请求)	复制源装置 ID (SD1(1))	复制源卷ID (α)	复制目的装置ID (SD1(2))	复制目的地卷 ID (β)

图19

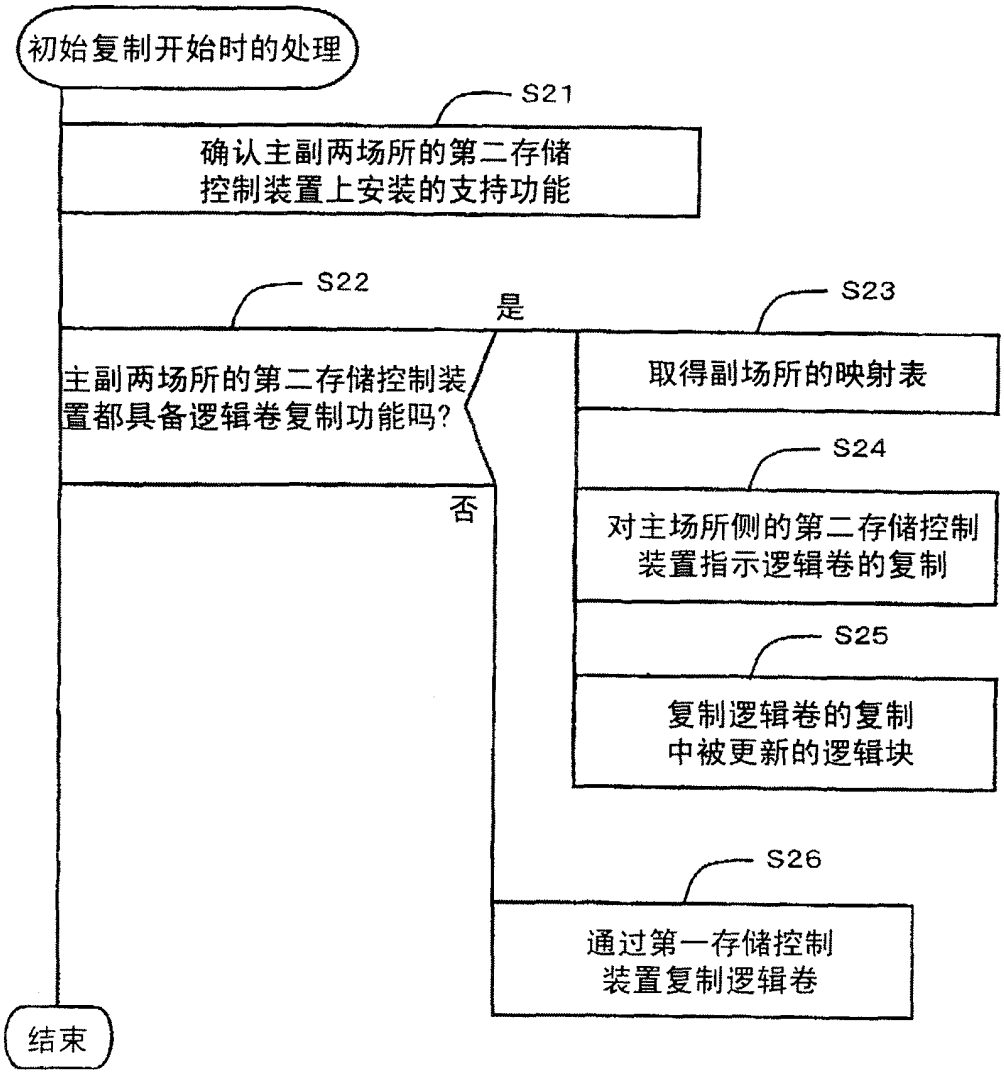

初始复制开始时的处理
S21
确认主副两场所的第二存储控制装置上安装的支持功能
S22
主副两场所的第二存储控制装置都具备逻辑卷复制功能吗?
是
否
S23
取得副场所的映射表
S24
对主场所侧的第二存储控制装置指示逻辑卷的复制
S25
复制逻辑卷的复制中被更新的逻辑块
S26
通过第一存储控制装置复制逻辑卷
结束

图20

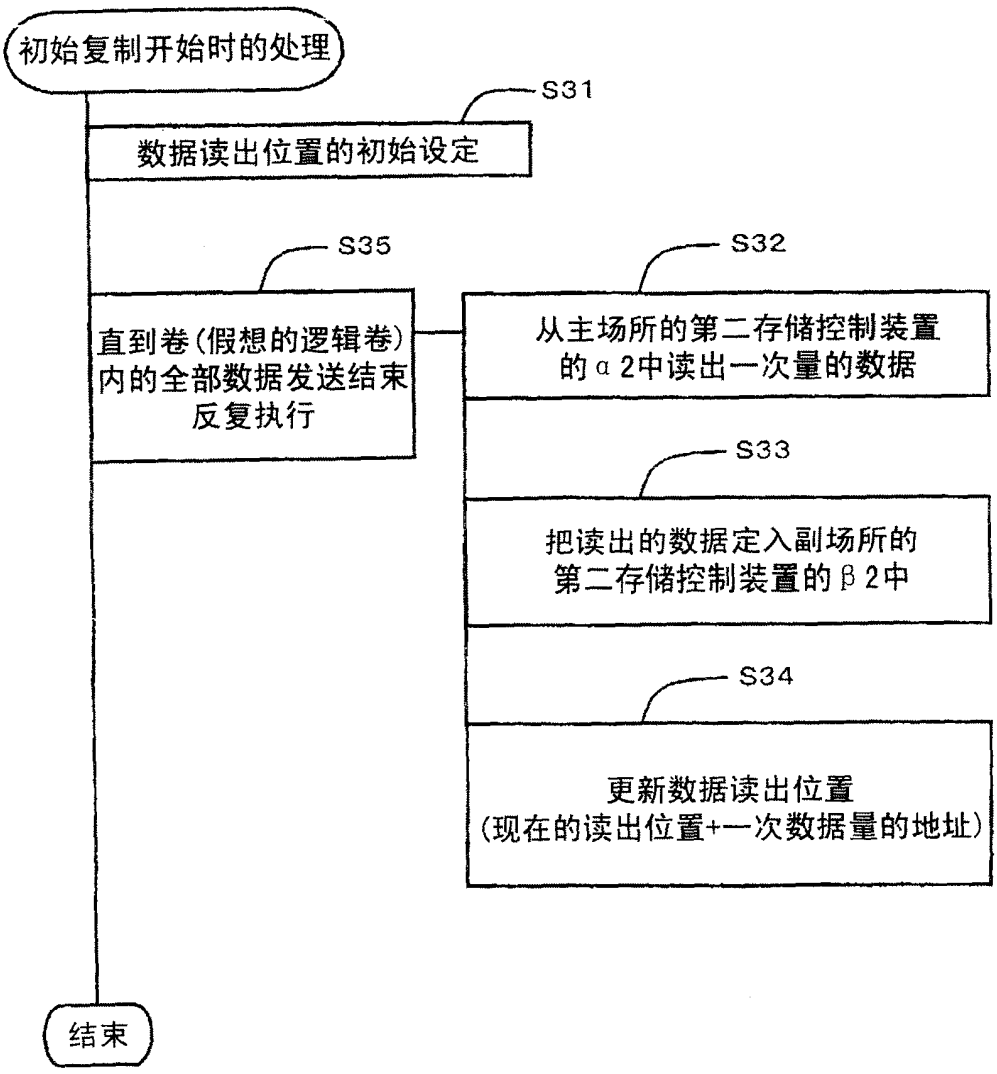

初始复制开始时的处理
S31
数据读出位置的初始设定
S35
直到卷(假想的逻辑卷)内的全部数据发送结束反复执行
S32
从主场所的第二存储控制装置的α2中读出一次量的数据
S33
把读出的数据定入副场所的第二存储控制装置的β2中
S34
更新数据读出位置(现在的读出位置+一次数据量的地址)
结束

图21

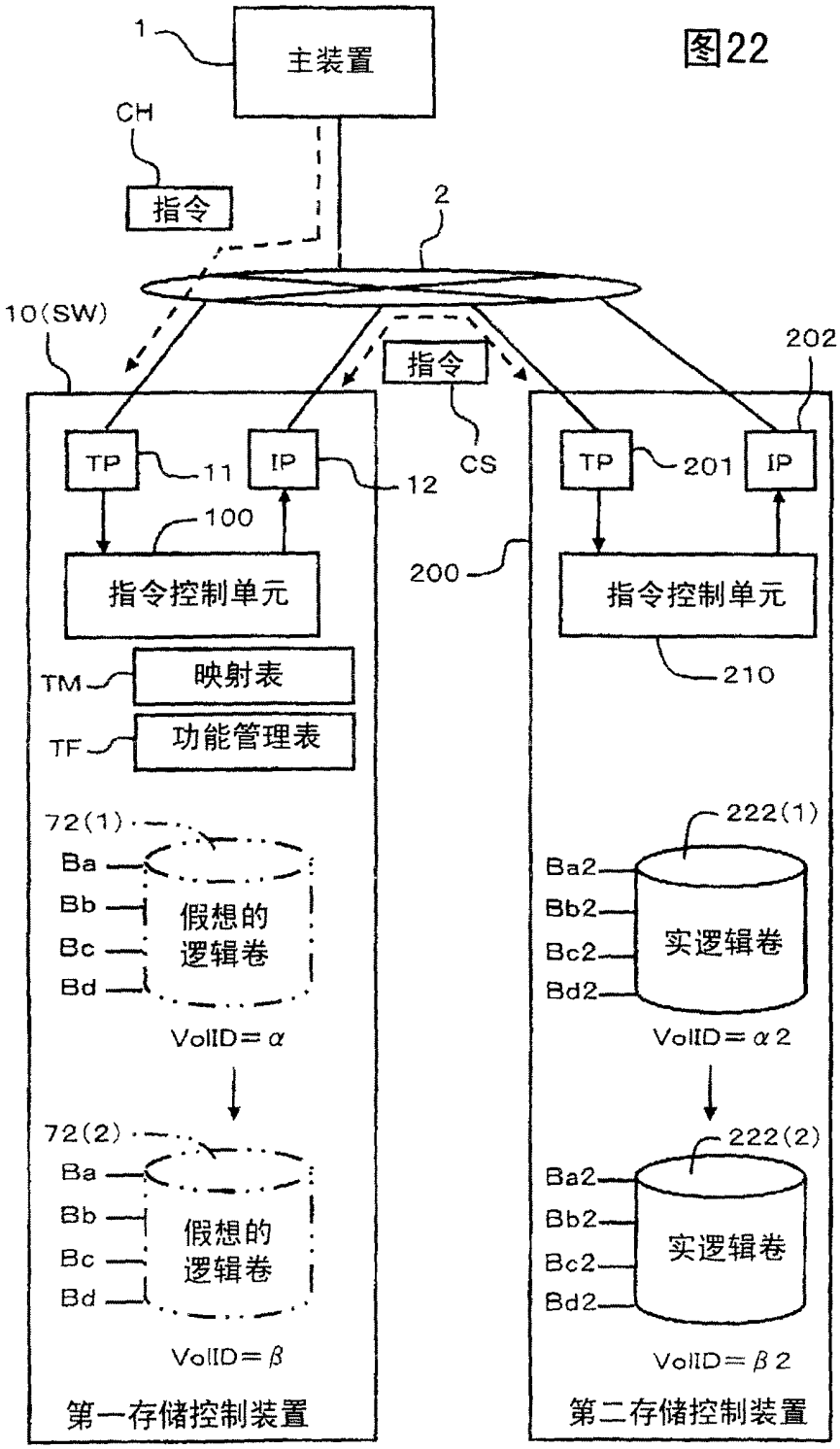

1
主装置
CH
指令
2
10(SW)
202
指令
CS
TP
11
IP
12
TP
201
IP
100
200
指令控制单元
指令控制单元
210
TM
映射表
TF
功能管理表
72(1)
Ba
Bb
Bc
Bd
假想的
逻辑卷
VolID=α
222(1)
Ba2
Bb2
Bc2
Bd2
实逻辑卷
VolID=α2
72(2)
Ba
Bb
Bc
Bd
假想的
逻辑卷
VolID=β
222(2)
Ba2
Bb2
Bc2
Bd2
实逻辑卷
VolID=β2
第一存储控制装置
第二存储控制装置

图22