



(12)发明专利

(10)授权公告号 CN 103518193 B

(45)授权公告日 2016.10.12

(21)申请号 201280023074.6

(72)发明人 约翰·C·吕德利克

(22)申请日 2012.03.27

(74)专利代理机构 北京律盟知识产权代理有限公司 11287

(65)同一申请的已公布的文献号  
申请公布号 CN 103518193 A

代理人 孙宝成

(43)申请公布日 2014.01.15

(51)Int. Cl.

(30)优先权数据

G06F 12/00(2006.01)

13/083,269 2011.04.08 US

G11C 7/10(2006.01)

G06F 13/14(2006.01)

(85)PCT国际申请进入国家阶段日  
2013.11.12

(56)对比文件

US 7908436 B1,2011.03.15,

(86)PCT国际申请的申请数据  
PCT/US2012/030713 2012.03.27

CN 101479944 A,2009.07.08,

CN 101286156 A,2008.10.15,

(87)PCT国际申请的公布数据  
W02012/138504 EN 2012.10.11

US 7747584 B1,2010.06.29,

审查员 李艳军

(73)专利权人 美光科技公司  
地址 美国爱达荷州

权利要求书5页 说明书6页 附图3页

(54)发明名称

用于数据去重复的方法及装置

(57)摘要

本发明包含用于数据去重复的装置及方法。一个此种方法包含:接收写入命令、变换与所述写入命令相关联的数据、确定所述数据的变换值是否存在于变换表中,及响应于所述变换值不存在于所述变换表中的确定,将与所述写入命令相关联的所述数据写入到存储器装置。

314

320	逻辑地址	散列 330	物理地址	340
320-1	LA-1	H-1 330-1	PA-1	340-1
320-2	LA-2	H-2 330-2	PA-2	340-2
320-3	LA-3	H-3 330-3	PA-3	340-3
320-4	LA-4	H-3 330-3	PA-3	340-3
320-5	LA-5	H-2 330-2	PA-2	340-2
320-6	LA-6	H-3 330-3	PA-3	340-3
320-M	LA-M	H-M 330-M	PA-M	340-M

1. 一种用于数据去重复的方法,其包括:

在存储器装置上的控制器处接收写入命令,其中所述控制器将所述写入命令翻译成若干个存储器装置写入命令,所述存储器装置写入命令的每一者都包含指令,当由所述控制器执行所述指令时,致使所述存储器装置进行将特定数量的数据写入到所述存储器装置中的写入操作;

通过创建所述数据的第一部分的第一变换值及与所述写入命令相关联的数据的第二部分的第二变换值,经由所述存储器装置的所述控制器上的数据去重复组件,变换与所述写入命令相关联的数据,其中所述数据的第一部分及所述数据的第二部分的每一者是比较经由所述写入操作写入的数据的所述特定数量少的数量的数据;

确定所述数据的所述第一部分所述第一变换值(330)是否存在于变换表(114、214、314)中;

确定所述数据的所述第二部分的所述第二变换值(330)是否存在于所述变换表中;及

响应于所述第一变换值(330)不存在于所述变换表(114、214、314)中,及所述第二变换值存在于所述变换表中的确定,通过将数据的所述第一部分写入到存储器装置(110、210)而跳过写入数据的所述第二部分,来执行所述写入操作,其中执行所述写入操作包括仅写入包括具有不在所述变换表中的变换值的数据部分的所述特定数量的数据;及

通过将数据的所述第一部分的逻辑地址与所述存储器装置中的数据所述第一部分的物理地址相关联,以及将数据的所述第二部分的逻辑地址与所述存储器装置中的数据所述第二部分的物理地址相关联来更新散列表,其中在所述写入操作的执行期间,数据的所述第一部分被写入到位于数据的所述第一部分的所述物理地址的存储器单元中,其中在先前执行的写入操作期间,数据的所述第二部分被写入到位于数据的所述第二部分的所述物理地址的存储器单元中。

2. 根据权利要求1所述的方法,其中接收所述写入命令包含从主机(102、202)接收所述写入命令。

3. 根据权利要求1所述的方法,其中确定所述数据的变换值(330)是否存在于变换表(114、214、314)中包含确定与所述写入命令相关联的所述数据的变换值(330)是否存在于存储于所述存储器装置(210)中的变换表(214、314)中。

4. 根据权利要求1所述的方法,其中确定所述数据的变换值(330)是否存在于变换表(114、214、314)中包含确定与所述写入命令相关联的所述数据的变换值(330)是否存在于存储于所述存储器装置(110)的控制器(108)中的变换表(114、314)中。

5. 根据权利要求1到4中任一权利要求所述的方法,其中所述数据包含若干个数据部分,且所述方法进一步包含将所述数据的一部分的所述变换值(330)写入到所述变换表(114、214、314)。

6. 根据权利要求1到4中任一权利要求所述的方法,其中所述数据包含若干个数据部分,且所述方法进一步包含经由所述变换表(114、214、314)使所述变换值(330)与所述数据的一部分的物理地址(340)相关联在一起。

7. 根据权利要求1到4中任一权利要求所述的方法,其中所述数据包含若干个数据部分,且变换与所述写入命令相关联的数据包含散列与所述写入命令相关联的所述数据的每一部分。

8. 一种用于数据去重复的方法,其包括:

经由控制器上的数据去重复组件散列与写入命令相关联的数据,其中所述控制器位于存储器装置上,且所述控制器将所述写入命令翻译成若干存储器装置写入命令,所述存储器装置写入命令的每一者都包含指令,当由所述控制器执行所述指令时,致使所述存储器装置进行将特定数量的数据写入到所述存储器装置中的写入操作,所述数据包括若干数据部分,所述数据部分的每一者包含比数据的第二数量少的第一数量的数据,其中数据的所述第二数量为经由所述写入操作写入的数据的所述特定数量,且散列与所述写入命令相关联的数据包括散列与所述写入命令相关联的每一个数据部分;

针对所述若干数据部分中的每一个确定与所述写入命令相关联的所述数据的散列值(330)是否存在于散列表(114、214、314)中;及

通过将与所述写入命令相关联的不具有所述散列表中的散列值的所述若干数据部分中的每一个写入到所述存储器装置(110、210),同时跳过写入散列值被确定为在所述散列表中的数据部分,来执行写入操作,其中所述写入操作包括仅写入具有不在所述散列表中的散列值的数据部分;及

将与所述写入命令相关联的不具有所述散列表中的散列值的所述若干数据部分中的每一个的逻辑地址与存储所述第一数量的数据并在所述写入操作的执行期间被写入的所述存储器装置中的存储器单元的物理地址相关联,以及将散列值被确定为在所述散列表中的每一数据部分的逻辑地址与存储所述第一数量的数据并在先前执行的写入操作期间被写入的所述存储器装置中的存储器单元的物理地址相关联,来更新所述散列表。

9. 根据权利要求8所述的方法,其中所述数据包含若干个数据部分,且散列与所述写入命令相关联的数据包含散列与所述写入命令相关联的数据的每一部分。

10. 根据权利要求9所述的方法,其包含将与所述写入命令相关联的所述数据的每一部分的不在所述散列表(114、214、314)中的所述散列值(330)写入到所述散列表(114、214、314)。

11. 根据权利要求10所述的方法,其中将与所述写入命令相关联的所述数据的每一部分的不在所述散列表(114、214、314)中的所述散列值(330)写入到所述散列表(114、214、314)包含写入到相变存储器单元。

12. 根据权利要求8到11中任一权利要求所述的方法,其进一步包含:如果所述数据的所述散列值(330)存在于所述散列表(114、214、314)中,那么将成功写入指示发送到主机(102、202),且使所述数据的逻辑地址(320)与存在于所述散列表(114、214、314)中的所述散列值相关联。

13. 根据权利要求8到11中任一权利要求所述的方法,其中确定与所述写入命令相关联的所述数据的散列值(330)是否存在于散列表(114、214、314)中包含经由所述存储器装置的控制器(108)的数据去重复电路(112)确定与所述写入命令相关联的所述数据的所述散列值(330)是否存在于所述散列表(114、214、314)中。

14. 一种用于数据去重复的方法,其包括:

向存储器装置发送写入命令以供所述存储器装置执行,其中控制器将所述写入命令翻译成若干个存储器装置写入命令,所述存储器装置写入命令的每一者都包含指令,当由所述控制器执行所述指令时,致使所述存储器装置进行将特定数量的数据写入到所述存储器

装置中的写入操作；

经由所述存储器装置上的数据去重复组件散列与所述写入命令相关联的数据，其中所述数据包含若干个数据部分，所述若干个数据部分的每一者包含比数据的第二数量少的第一数量的数据，其中数据的所述第二数量为经由所述写入操作写入的数据的所述特定数量；

确定与所述写入命令相关联的数据的每一部分的散列值(330)是否存在于散列表(114、214、314)中；

响应于散列值(330)存在于所述散列表(114、214、314)中的确定，针对与存在于所述散列表中的散列值相关联的数据部分，将成功写入指示发送到主机(102、202)；及

响应于散列值(330)不存在于所述散列表(114、214、314)中的确定，通过将所述写入命令相关联的数据写入到所述存储器装置(110、210)，同时跳过写入散列值被确定为在所述散列表中的数据部分，来执行写入操作，其中写入数据包括仅写入具有不在所述散列表中的散列值的数据部分；及

通过将所述写入命令相关联的不具有所述散列表中的散列值的所述若干数据部分中的每一个的逻辑地址与存储所述第一数量的数据并在所述写入操作的执行期间被写入的所述存储器装置中的存储器单元的物理地址相关联，以及将散列值被确定为在所述散列表中的每一数据部分的逻辑地址与存储所述第一数量的数据并在先前执行的写入操作期间被写入的所述存储器装置中的存储器单元的物理地址相关联，来更新所述散列表。

15. 根据权利要求14所述的方法，所述方法进一步包含使所述数据的每一部分的逻辑地址(320)与所述数据的每一部分的存在于所述散列表(114、214、314)中的所述散列值(330)相关联。

16. 根据权利要求14所述的方法，所述方法进一步包含经由所述散列表(114、214、314)使所述数据的每一部分的逻辑地址(320)与所述数据的每一部分的所述散列值(330)相关联且与所述数据的每一部分的物理地址(340)相关联。

17. 根据权利要求14到16中任一权利要求所述的方法，其进一步包含将每一部分的不在所述散列表(114、214、314)中的所述散列值(330)写入到所述散列表(114、214、314)。

18. 根据权利要求14到16中任一权利要求所述的方法，其进一步包含经由所述散列表(114、214、314)使共享共同散列值(330)的若干个逻辑地址(320)与所述存储器装置(110、210)中的共同物理位置(340)相关联。

19. 一种用于数据去重复的装置，其包括：

存储器单元阵列(110、210)；及

控制器(108、208)，其耦合到所述阵列且经配置以：

接收写入命令，并将所述写入命令翻译成若干个存储器装置写入命令，所述存储器装置写入命令的每一者都包含指令，当由所述控制器执行所述指令时，致使所述存储器单元阵列进行将特定数量的数据写入到所述存储器单元阵列中的写入操作；

散列与所述写入命令相关联的若干个数据部分的每一部分，所述若干个数据部分的每一部分包含比数据的第二数量少的第一数量的数据，其中数据的所述第二数量为经由所述写入操作写入的数据的所述特定数量；

经由所述控制器上的数据去重复组件，确定与所述写入命令相关联的所述若干个数据

部分的每一者的散列值(330)是否存在于所述存储器单元阵列的散列表(114、214、314)中；及

仅在确定所述散列值(330)不存在于所述散列表(114、214、314)中的情况下,通过将对应于不存在于所述散列表(114、214、314)中的所述散列值(330)的与所述写入命令相关联的所述若干个数据部分的每一者的写入到所述存储器单元阵列(110、210),同时跳过写入散列值被确定为不在所述散列表中的数据部分,来执行写入操作,其中所述写入操作包括仅写入具有不在所述散列表中的散列值的数据部分;及

通过将所述写入命令相关联的不具有所述散列表中的散列值的所述若干个数据部分中的每一个的逻辑地址与存储所述第一数量的数据并在所述写入操作的执行期间被写入的所述存储器装置中的存储器单元的物理地址相关联,以及将散列值被确定为在所述散列表中的每一数据部分的逻辑地址与存储所述第一数量的数据并在先前执行的写入操作期间被写入的所述存储器装置中的存储器单元的物理地址相关联,来更新所述散列表。

20. 根据权利要求19所述的装置,其中所述控制器(108)在包含所述阵列(110)的存储器装置的外部。

21. 根据权利要求19所述的装置,其中所述控制器(208)在包含所述阵列(210)的存储器装置的内部。

22. 根据权利要求19所述的装置,其中所述散列表(114、314)存储于所述控制器(108)中。

23. 根据权利要求19所述的装置,其中所述散列表(214、314)存储于所述阵列(210)中。

24. 根据权利要求19所述的装置,其中所述数据包含若干个数据部分,且所述控制器经配置以响应于所述散列值(330)存在于所述散列表(114、214、314)中的确定而将成功写入指示发送到主机(102、202)且使数据的一部分的逻辑地址与存在于所述散列表(114、214、314)中的所述散列值(330)相关联。

25. 根据权利要求19到24中任一权利要求所述的装置,其中所述数据包含若干个数据部分,且所述散列表(114、214、314)经由散列值使数据的一部分的逻辑地址(320)与数据的所述部分的物理地址(340)相关联。

26. 一种用于数据去重复的装置,其包括:

存储器单元阵列(110、210);及

控制器(108、208),其耦合到所述阵列且经配置以:

散列与从主机(102、202)接收的写入命令相关联的数据,其中所述控制器将所述写入命令翻译成若干个存储器装置写入命令,所述存储器装置写入命令的每一者都包含指令,当由所述控制器执行所述指令时,致使所述存储器单元阵列进行将特定数量的数据写入到所述存储器单元阵列中的写入操作,其中所述数据包含若干个数据部分,且其中所述若干个数据部分的每一者是比较经由所述写入操作写入的数据的所述特定数量少的第一数量的数据;

经由所述控制器上的数据去重复组件,确定与所述写入命令相关联的所述若干个数据部分的每一者的散列值(330)是否存在于所述存储器单元阵列中的散列表(114、214、314)中;

如果数据的一部分的散列值(330)存在于所述散列表(114、214、314)中,那么针对数据

的所述部分将成功写入指示发送到所述主机(102、202);及

如果数据的所述部分的所述散列值(330)不在所述散列表(114、214、314)中,那么将数据的所述部分写入到所述存储器单元阵列,同时跳过写入其散列值包含在所述散列表中的数据部分,其中所述写入操作包括仅写入具有不在所述散列表中的散列值的数据部分;及

通过将所述写入命令相关联的不具有所述散列表中的散列值的所述若干个数据部分中的每一个的逻辑地址与存储比数据的所述特定数量少的数量的数据并在所述写入操作的执行期间被写入的所述存储器装置中的存储器单元的物理地址相关联,以及将散列值被确定为在所述散列表中的每一数据部分的逻辑地址与存储比数据的所述特定数量少的数据并在先前执行的写入操作期间被写入的所述存储器装置中的存储器单元的物理地址相关联,来更新所述散列表。

27. 根据权利要求26所述的装置,其中经配置以确定与所述写入命令相关联的数据的每一部分的所述散列值(330)是否存在于所述散列表(114、214、314)中的所述控制器包括数据去重复电路(112、212)。

28. 根据权利要求26到27中任一权利要求所述的装置,其中所述控制器经配置以:如果确定数据的每一部分的所述散列值不存在于所述散列表(114、214、314)中,那么将所述散列值(330)写入到所述散列表(114、214、314)。

29. 一种用于数据去重复的装置,其包括:

存储器单元阵列(110、210);及

控制器(108、208),其耦合到所述阵列且经配置以:

接收供所述装置执行的写入命令,其中所述控制器将所述写入命令翻译成若干个存储器装置写入命令,所述存储器装置写入命令的每一者都包含指令,当由所述装置执行所述指令时,致使所述存储器单元阵列进行将特定数量的数据写入到所述存储器单元阵列中的写入操作

经由所述控制器上的数据去重复组件,散列若干个数据部分中的每一者,所述若干个数据部分中的每一者包含比数据的第二数量少的第一数量的数据,其中数据的所述第二数量为经由所述写入操作写入的数据的所述特定数量;

针对所述若干个部分中的每一部分,确定与所述部分相关联的散列值(330)是否存在于所述存储器单元阵列中的散列表(114、214、314)中;及

仅针对其散列值(330)经确定不存在于所述散列表(114、214、314)中的那些数据部分,通过将对应于那些数据部分的数据写入于所述阵列(110、210)中,同时跳过写入散列值被确定为在所述散列表中的数据部分,来执行所述写入操作,其中所述写入操作包括仅写入具有不在所述散列表中的散列值的数据部分;及

通过将所述写入命令相关联的不具有所述散列表中的散列值的所述若干个数据部分中的每一个的逻辑地址与存储所述第一数量的数据并在所述写入操作的执行期间被写入的所述存储器装置中的存储器单元的物理地址相关联,以及将散列值被确定为在所述散列表中的每一数据部分的逻辑地址与存储所述第一数量的数据并在先前执行的写入操作期间被写入的所述存储器装置中的存储器单元的物理地址相关联,来更新所述散列表。

## 用于数据去重复的方法及装置

### 技术领域

[0001] 本发明大体来说涉及半导体存储器装置及方法,且更特定来说,涉及用于数据去重复的装置及方法。

### 背景技术

[0002] 通常提供存储器装置作为计算机或其它电子装置中的内部半导体集成电路及/或外部可装卸装置。存在许多不同类型的存储器,包含随机存取存储器(RAM)、只读存储器(ROM)、动态随机存取存储器(DRAM)、同步动态随机存取存储器(SDRAM)、相变随机存取存储器(PCRAM)及快闪存储器以及其它存储器。

[0003] 利用存储器装置作为用于需要高存储器密度、高可靠度及低电力消耗的宽广范围的电子应用的非易失性存储器。非易失性存储器可用于个人计算机、便携式存储器棒、固态驱动器(SSD)、个人数字助理(PDA)、数码相机、蜂窝式电话、便携式音乐播放器(例如,MP3播放器)、电影播放器及其它电子装置以及其它装置中。例如程序代码的数据及系统数据(例如基本输入/输出系统(BIOS))通常存储于非易失性存储器装置中。

[0004] 存储器系统可从主机接收写入命令,其中所述写入命令与待写入到存储器系统的一个或一个以上存储器装置的数据相关联(例如,包含所述数据)。与写入命令相关联(例如,在所述写入命令中)的数据可包含若干个部分(例如,扇区),其中每一部分具有与其相关联的唯一逻辑地址。可给每一逻辑地址指派与物理位置相关联的物理地址,且可将写入命令中的数据的一部分写入到与指派给对应于所述数据的一部分的所述逻辑地址的物理地址相关联的特定物理位置。在各种实例中,不同逻辑地址可对应于相同物理数据使得所述数据在存储器系统中的一个或一个以上存储器装置中物理上重复。写入重复数据可不必要地减少存储器系统的存储容量以及其它缺陷。

### 发明内容

### 附图说明

[0005] 图1是根据本发明的一个或一个以上实施例的包含至少一个存储器系统的计算装置的功能框图。

[0006] 图2是根据本发明的一个或一个以上实施例的包含至少一个存储器系统的计算装置的功能框图。

[0007] 图3图解说明根据本发明的一个或一个以上实施例的表示若干个逻辑地址、散列值及物理地址的散列表。

### 具体实施方式

[0008] 本发明包含用于数据去重复的装置及方法。一个此种方法包含接收写入命令、变换与写入命令相关联(例如,来自写入命令)的数据、确定数据的变换值是否存在于变换表

中,及响应于变换值不存在于变换表中的确定,将与写入命令相关联的数据写入到存储器装置。

[0009] 数据去重复可包含避免将重复数据写入到存储器系统。在先前方法中,主机将待写入到存储器系统的文件与存储器系统上的文件进行比较以确定文件是否已存在于存储器系统上且删除存在于存储器系统上的现有重复文件。在这些方法中,如果文件存在于存储器系统上,那么主机将不把写入所述文件的命令发送到存储器系统,因此避免文件在存储器系统上的重复。在一个或一个以上实施例,数据去重复可包含变换(例如,散列)由存储器系统从主机接收的每一数据扇区。每一数据扇区的散列可包含产生针对在写入命令中接收的每一数据扇区的变换值(例如,散列值)。实施例不限于散列每一数据扇区,且举例来说,数据可以任何增量(例如页)来散列。

[0010] 在一个或一个以上实施例中,可将针对每一数据扇区而产生的散列值与变换表(例如,散列表)中的针对当前写入到存储器系统中的存储器装置的数据的散列值进行比较。如果针对在写入命令中接收的数据扇区而产生的散列值不存在于散列表中(例如,在散列表中未呈现),那么将数据扇区写入到存储器装置中的位置且通过将包含扇区数据的散列值的条目写入到散列表来更新所述散列表。还可通过将包含数据扇区的逻辑地址及数据扇区的物理地址的条目写入到散列表来更新所述散列表。如果针对在写入命令中接收的数据扇区而产生的散列值确实存在于散列表中,那么通过将使数据扇区的逻辑地址与扇区数据的散列值相关联的条目写入到散列表来更新所述散列表。在用数据扇区的散列值以及逻辑地址及物理地址更新散列表时,可将成功写入指示发送到主机。

[0011] 在本发明的以下详细描述中,参考形成本发明的一部分的附图,且其中以图解说明的方式展示可如何实践本发明的若干个实施例。充分详细地描述这些实施例以使所属领域的技术人员能够实践本发明的实施例,且应理解,可利用其它实施例且可做出过程、电及/或结构改变,此并不背离本发明的范围。

[0012] 如本文中所以使用,“若干个”某物可指代一个或一个以上此事物。举例来说,若干个存储器装置可指代一个或一个以上存储器装置。另外,如本文中所以使用的标示符“N”及“M”(特定来说关于图式中的参考编号)指示本发明的若干个实施例可包含若干个如此标示的特定特征。

[0013] 本文中的图遵循其中第一个数字或前几个数字对应于图式图编号且剩余数字识别图式中的元件或组件的编号惯例。不同图之间的类似元件或组件可通过使用类似数字来识别。举例来说,在图1中100可指代元件“00”,且在图2中类似元件可指代200。如将了解,可添加、交换及/或消除本文中的各种实施例中所展示的元件以便提供本发明的若干个额外实施例。另外,如将了解,图中所提供的元件的比例及相对比例尺既定图解说明本发明的实施例,且不应视为限定意义。

[0014] 图1是根据本发明的一个或一个以上实施例的包含至少一个存储器系统104的计算装置100的功能框图。在图1中所图解说明的实施例中,存储器系统104可包含控制器108及一个或一个以上存储器装置110-1、…、110-N。在此实例中,控制器108在一个或一个以上存储器装置110-1、…、110-N的外部。存储器装置110-1、…、110-N可提供用于存储器系统的存储卷,例如借助格式化到存储器装置的文件系统。控制器108可包含控制电路,例如硬件、固件及/或软件。在一个或一个以上实施例中,控制器108可为耦合到包含物理接口及存储



器装置110-1、…、110-N的印刷电路板的专用集成电路(ASIC)。

[0015] 控制器108可包含数据去重复组件112及散列表114。数据去重复组件112可包含硬件、固件及/或软件,且可对数据执行数据去重复操作,例如,以确定数据是否存在于存储器装置110-1、…、110-N中及是否应将数据写入到存储器装置110-1、…、110-N。散列表114可存储于控制器108上的存储器单元中及/或存储器装置110-1、…、110-N的存储器单元中,且可用于使逻辑地址与散列值相关联且也与数据的一部分(例如扇区)的物理地址相关联。散列表114可包含对应于数据的部分(例如,扇区)的逻辑地址、散列值及物理地址的若干个条目。数据去重复组件112可散列由存储器系统从主机接收的每一数据扇区。在散列每一数据扇区时,可产生在写入命令中接收的每一数据扇区的散列值。

[0016] 在一个或一个以上实施例中,数据去重复组件112可将针对每一数据扇区而产生的散列值与对应于散列表112中当前写入到存储器装置110-1、…、110-N的数据的散列值进行比较。如果针对在写入命令中接收的数据扇区而产生的散列值不存在于散列表114中,那么将所述数据扇区写入到存储器装置110-1、…、110-N中的位置且通过将具有所述数据扇区的散列值的条目写入到散列表114来更新散列表114。还可通过将使数据扇区的逻辑地址及数据扇区的物理地址与所述数据扇区的散列值相关联的条目写入到散列表114来更新散列表114。如果针对在写入命令中接收的数据扇区而产生的散列值确实存在于散列表114中,那么可通过将使数据扇区的逻辑地址与已经在散列表114中的扇区数据散列值相关联的条目写入到散列表114来更新散列表114。在用数据扇区的散列值以及逻辑地址及物理地址更新散列表114时,可将成功写入指示发送到主机102。

[0017] 如图1中所图解说明,主机102可耦合到存储器系统104。主机102可为膝上型计算机、个人计算机、数码相机、数字记录及回放装置、移动电话、PDA、存储器卡读取器、接口集线器以及其它主机系统,且可包含存储器存取装置,例如,处理器。所属领域的技术人员将了解,“处理器”可意指一个或一个以上处理器,例如并行处理系统、若干个协处理器等。

[0018] 在一个或一个以上实施例中,物理主机接口可呈标准化接口的形式。举例来说,当存储器系统104用于计算装置100中的数据存储时,物理主机接口可为串行高级技术附件(SATA)、高速外围组件互连(PCIe)或通用串行总线(USB)以及其它接口。然而,一般来说,物理主机接口可提供用于在存储器系统104与具有用于物理主机接口的兼容接受器的主机102之间传递控制、地址、数据及其它信号的接口。主机102可经由物理主机接口将写入命令发送到存储器系统104。

[0019] 控制器108可与存储器装置110-1、…、110-N通信以读取、写入及擦除数据,以及进行其它操作。控制器108可具有可为一个或一个以上集成电路及/或离散组件的电路。存储器控制器可选择性地耦合存储器装置110-1、…、110-N的I/O连接(未在图1中展示)以在适当的时间在适当的I/O连接处接收适当的信号。类似地,主机102与存储器系统104之间的通信协议可不同于存取存储器装置110-1、…、110-N所需的协议。接着,控制器108可将从主机接收的命令转译成适当的命令以实现对于存储器装置110-1、…、110-N的所要存取。

[0020] 存储器装置110-1、…、110-N可包含一个或一个以上存储器单元(例如,非易失性存储器单元)阵列。举例来说,所述阵列可为相变存储器(PCM)阵列。实施例不限于特定类型的存储器装置。例如,存储器装置可包含RAM、ROM、HDD、DRAM、SDRAM、PCRAM、RRAM及快闪存储器及其它存储器。

[0021] 存储器装置110-1、…、110-N可包含可分组的若干个存储器单元。如在本文中所使用,一群组可包含一个或一个以上存储器单元,例如,页、块、平面、裸片、整个阵列或其它存储器单元群组。举例来说,一些存储器阵列可包含构成存储器单元块的若干个存储器单元页。若干个块可包含于存储器单元平面中。若干个存储器单元平面可包含于裸片上。作为一实例,128GB存储器装置可包含每页4314个数据字节,每块128个页,每平面2048个块及每装置16个平面。

[0022] 图1的实施例可包含为不使本发明的实施例模糊而未图解说明的额外电路。举例来说,存储器系统104可包含控制器(例如,控制电路)及地址电路以锁存通过I/O电路经由I/O连接提供的地址信号。可由行解码器及列解码器接收并解码地址信号以存取存储器装置110-1、…、110-N。所属领域的技术人员将了解,地址输入连接的数目可取决于存储器装置110-1、…、110-N的密度及架构。

[0023] 一般来说,控制器108可负责将从主机102(例如,从PCIe总线)接收的命令转换成用于主机-存储器转译电路的命令指令且负责将存储器响应转换成用于传输到发请求的主机的主机系统命令。

[0024] 在一个或一个以上实施例中,可将数据一次一页地写入到存储器装置。存储器装置中的每一页可具有若干个物理扇区且每一物理扇区可与一LBA相关联。作为一实例,物理页可具有8个物理数据扇区。然而,实施例不限于每物理页特定数目个物理扇区。

[0025] 图2是根据本发明的一个或一个以上实施例的包含至少一个存储器系统204的计算装置200的功能框图。在图2中所图解说明的实施例中,存储器系统204可包含控制器208及一个或一个以上存储器装置210-1、…、210-N。在此实例中,控制器208在一个或一个以上存储器装置210-1、…、210-N的外部,但在其它实施例中,所述控制器可在所述存储器装置中的一者或一者以上的内部或分布在所述存储器装置中的一者或一者以上及/或控制器208之间。存储器装置210-1、…、210-N可提供用于存储器系统的存储卷,例如,借助格式化到存储器装置的文件系统。控制器208可包含控制电路,例如硬件、固件及/或软件。在一个或一个以上实施例中,控制器208可为耦合到包含物理接口及存储器装置210-1、…、210-N的印刷电路板的专用集成电路(ASIC)。

[0026] 存储器装置210-1、…、210-N可包含数据去重复组件212,例如内部控制器(例如,控制电路)的一部分。数据去重复组件212可包含硬件、固件及/或软件,且可对数据执行数据去重复操作,如以上结合图1所论述。存储器装置210-1、…、210-N可包含散列表214。所述散列表可存储于存储器装置210-1、…、210-N的存储器单元中,且可用于使逻辑地址与散列值相关联且也与数据的一部分(例如扇区)的物理地址相关联,如以上结合图1所论述。

[0027] 控制器208可与存储器装置210-1、…、210-N通信以读取、写入及擦除数据,以及进行其它操作。控制器208可具有可为一个或一个以上集成电路及/或离散组件的电路。存储器控制器可选择性地耦合存储器装置210-1、…、210-N的I/O连接(未在图1中展示)以在适当的时间在适当的I/O连接处接收适当的信号。类似地,主机202与存储器系统204之间的通信协议可不同于存取存储器装置210-1、…、210-N所需的协议。由此控制器208可将主机接收的命令转译成适当的命令以实现存取存储器装置210-1、…、210-N的所要存取。

[0028] 存储器装置210-1、…、210-N可包含一个或一个以上存储器单元(例如,非易失性存储器单元)阵列。举例来说,所述阵列可为相变存储器(PCM)阵列。实施例不限于特定类型

的存储器装置。例如,存储器装置可包含RAM、ROM、DRAM、SDRAM、PCRAM、RRAM及快闪存储器以及其它存储器。

[0029] 存储器装置210-1、...、210-N可包含可分组的若干个存储器单元。如本文中所使用,一群组可包含一个或一个以上存储器单元,例如,页、块、平面、裸片、整个阵列或其它存储器单元群组。举例来说,一些存储器阵列可包含构成存储器单元块的若干个存储器单元页。若干个块可包含于存储器单元平面中。若干个存储器单元平面可包含于裸片上。作为一实例,128GB存储器装置可包含每页4314个数据字节,每块128个页,每平面2048个块及每装置16个平面。

[0030] 图3图解说明根据本发明的一个或一个以上实施例的表示若干个逻辑地址、散列值及物理地址的散列表314。图3中的散列表314包含若干个列。举例来说,第一列320指示特定数据部分(例如扇区)的逻辑地址(LA)。第二列330指示特定数据部分的散列值。第三列340指示特定数据部分的物理地址(PA)。特定数据部分的物理地址包含指向存储器装置中的所述数据部分的物理位置的指针。

[0031] 在图3中所展示的实例中,列320包含若干个不同逻辑地址320-1(“LA-1”)、320-2(“LA-2”)、320-3(“LA-3”)、320-4(“LA-4”)、320-5(“LA-5”)、320-6(“LA-6”)及320-M(“LA-M”)。每一逻辑地址与相应散列值330-1(“H-1”)、330-2(“H-2”)、330-3(“H-3”)及330-M(“H-M”)相关联。

[0032] 在一个或一个以上实施例中,散列表314的列320中的逻辑地址及其在列330中的相关联散列值可与指示存储器系统(例如,图1中的系统104)中的数据的物理位置的列340中的物理地址相关联、与列320中的逻辑地址及列330中的散列值相关联。每一逻辑地址320-1、320-2、320-3、320-4、320-5、320-6及320-M以及相关散列值330-1、330-2、330-3及330-M可具有相关联物理地址340-1(“PA-1”)、340-2(“PA-2”)、340-3(“PA-3”)及340-M(“PA-M”)。共享共同散列值的逻辑地址可与相同物理位置相关联。此允许将与两个或两个以上逻辑地址相关联的数据仅写入到存储器装置一次。

[0033] 在图3中所展示的实例中,逻辑地址320-1与散列值330-1相关联且与逻辑地址320-1相关联的数据在存储器装置中的与物理地址340-1相关联的物理位置处。逻辑地址320-2及逻辑地址320-5与散列值330-2相关联且与逻辑地址320-2及逻辑地址320-5相关联的数据在存储器装置中的与物理地址340-2相关联的物理位置处。

[0034] 逻辑地址320-3、逻辑地址320-4及逻辑地址320-6与散列值330-3相关联且与逻辑地址320-3、逻辑地址320-4及逻辑地址320-6相关联的数据在存储器装置中的与物理地址340-3相关联的物理位置处。逻辑地址320-M与散列值330-M相关联且与逻辑地址320-M相关联的数据在存储器装置中的与物理地址340-M相关联的物理位置处。

[0035] 逻辑地址320-1与散列值330-1相关联且与逻辑地址320-1相关联的数据在存储器装置中的与物理地址340-1相关联的物理位置处。在一些实施例中,与由存储器系统接收的写入命令相关联(例如,为所述写入命令的一部分)的共享共同散列值(例如,散列值330-1)的任何数据将不写入到存储器装置,但可使其逻辑地址与共同散列值330-1及共同物理地址340-1相关联。

[0036] 在此实例中,逻辑地址320-2及320-5对应于相同数据,因此其共享共同散列值(散列值330-2)。将与逻辑地址320-2及320-5相关联的数据写入到与物理地址340-2相关联的

物理位置一次,因此将与逻辑地址320-2及320-5相关联的数据仅写入到存储器装置一次。在一些实施例中,与由存储器系统接收的写入命令相关联的共享共同散列值(例如,散列值330-2)的任何数据将与具有共同物理地址(例如,340-2)的共同物理位置相关联。当存取(例如,在读取命令中请求)与逻辑地址320-2及320-5相关联的数据时,可使用散列表314经由与逻辑地址320-2及320-5相关联的物理地址340-2来识别数据的物理位置,且可从相同物理位置读取与逻辑地址320-2及320-5相关联的读取命令。

[0037] 在此实例中,逻辑地址320-3、320-4及320-6对应于相同数据,因此其共享共同散列值(散列值330-3)。将与逻辑地址320-3、320-4及320-6相关联的数据写入到与物理地址340-3相关联的物理位置一次,因此将与逻辑地址320-3、320-4及320-6相关联的数据仅写入到存储器装置一次。在一些实施例中,与由存储器系统接收的写入命令相关联的共享共同散列值(例如,散列值330-3)的任何数据将与具有共同物理地址(例如,340-3)的共同物理位置相关联。当存取(例如,在读取命令中请求)与逻辑地址320-3、320-4及320-6相关联的数据时,可使用散列表314经由与逻辑地址320-3、320-4及320-6相关联的物理地址340-3识别数据的物理位置,且从相同物理位置读取与逻辑地址320-3、320-4及320-6相关联的读取命令。

[0038] 在一个或一个以上实施例中,散列表可包含散列值及与散列值相关联的数据的物理位置。在一些实施例中,与存在于散列表中的散列值相关联的数据将不写入到存储器装置。散列表可包含散列值的列表及与每一散列值相关联的物理位置的列表,其可用于确定与散列值相关联的数据是否位于存储器装置中的物理位置处及是否不必将与散列值相关联的数据写入于存储器装置中(因为散列值已位于存储器装置中的物理位置处)。

[0039] 结论

[0040] 本发明包含用于数据去重复的装置及方法。一个此种方法包含:接收写入命令、变换与写入命令相关联的数据、确定数据的变换值是否存在于变换表中,及响应于变换值不存在于变换表中的确定,将与写入命令相关联的数据写入到存储器装置。

[0041] 尽管本文中已图解说明且描述了特定实施例,但所属领域的技术人员将了解,可用旨在实现相同结果的布置来替代所展示的特定实施例。本发明既定涵盖本发明的若干个实施例的改动或变化形式。应理解,以上描述是以说明性方式而非限制性方式作出。在审阅以上描述后,所属领域的技术人员将明了以上实施例的组合及未在本文中具体描述的其它实施例。本发明的若干个实施例的范围包含其中使用以上结构及方法的其它应用。因此,应参考所附权利要求书范围连同授权此权利要求书的等效内容的全部范围一起来确定本发明的若干个实施例的范围。

[0042] 在前述具体实施方式中,出于简化本发明的目的而将一些特征一同集合于单个实施例中。不应将本发明的此方法解释为反映本发明的所揭示实施例必须使用比每一权利要求中明确陈述的特征多的特征的意图。而是,如所附权利要求书反映,发明性标的物在于少于单个所揭示实施例的所有特征。因此,特此将所附权利要求书并入到具体实施方式中,其中每一权利要求独立地作为单独实施例。

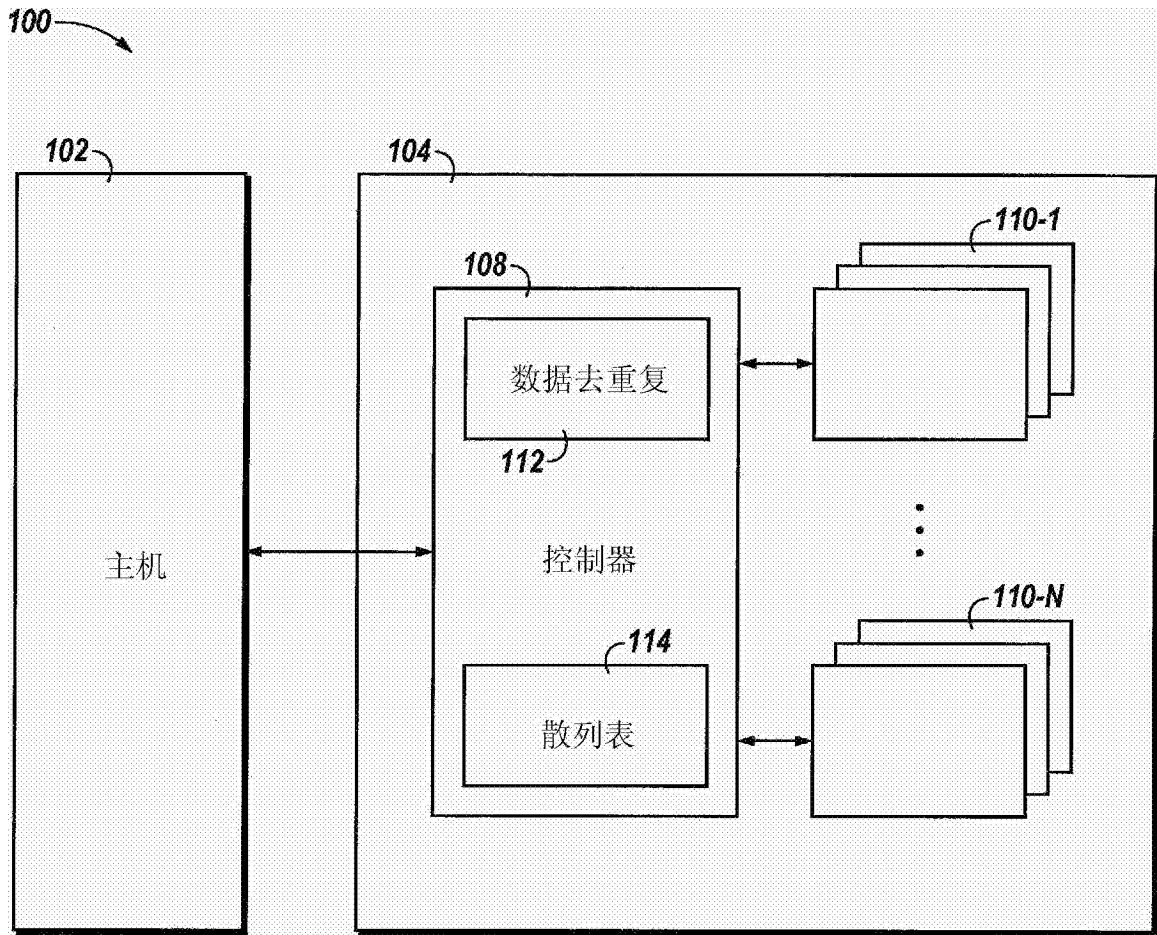


图1

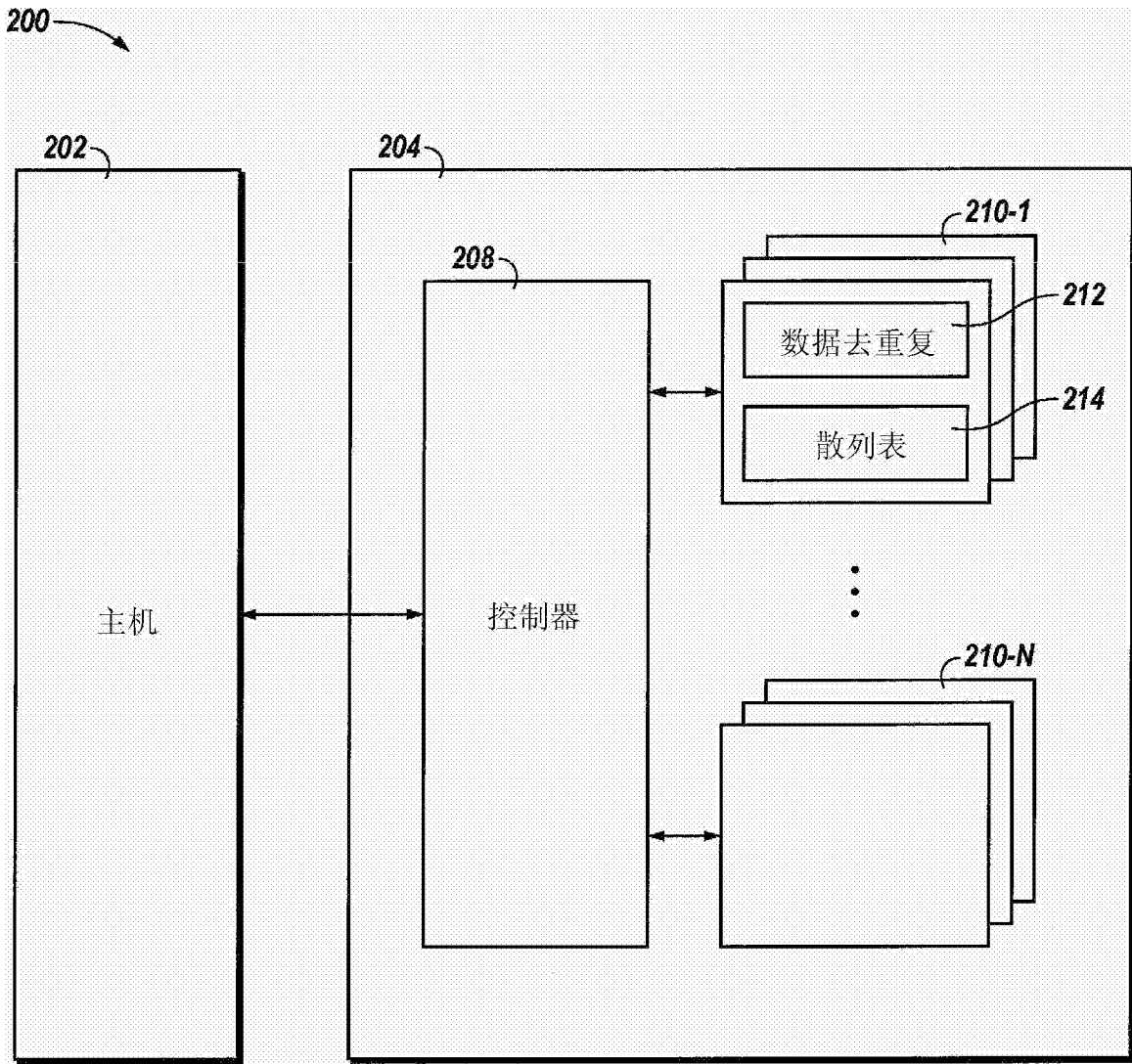


图2

314

320	逻辑地址	散列 <u>330</u>	物理地址	340
320-1	LA-1	H-1 <u>330-1</u>	PA-1	340-1
320-2	LA-2	H-2 <u>330-2</u>	PA-2	340-2
320-3	LA-3	H-3 <u>330-3</u>	PA-3	340-3
320-4	LA-4	H-3 <u>330-3</u>	PA-3	340-3
320-5	LA-5	H-2 <u>330-2</u>	PA-2	340-2
320-6	LA-6	H-3 <u>330-3</u>	PA-3	340-3
320-M	LA-M	H-M <u>330-M</u>	PA-M	340-M

图3