(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101512491 B
(45)授权公告日 2012.11.28

(21)申请号 200680020980.5

(22)申请日 2006.04.17

(30)优先权数据
11/107,396 2005.04.15 US

(85)PCT申请进入国家阶段日
2007.12.12

(86)PCT申请的申请数据
PCT/US2006/014772 2006.04.17

(87)PCT申请的公布数据
WO2006/113823 EN 2006.10.26

(73)专利权人 英特尔公司
地址 美国加利福尼亚州

(72)发明人 F·科拉多 D·内米罗夫

(74)专利代理机构 中国专利代理(香港)有限公司 72001
代理人 曾祥夌 刘春元

(51)Int.Cl.
G06F 11/10(2006.01)

(56)对比文件
CN 1503224 A,2004.06.09, 全文.
US 6067635 A,2000.05.23, 全文.
US 5911779 A,1999.06.15, 说明书第11栏第12至22行.
US 5774643 A,1998.06.30, 权利要求18、说明书第4栏第50至53行,第11栏第1至20行,第5栏第44至65行,第9栏第40至第10栏65行,第12栏第25行至第14栏第55行、附图2.
US 5774643 A,1998.06.30, 权利要求18、说明书第4栏第50至53行,第11栏第1至20行,第5栏第44至65行,第9栏第40至第10栏65行,第12栏第25行至第14栏第55行、附图2.

审查员 刘长勇

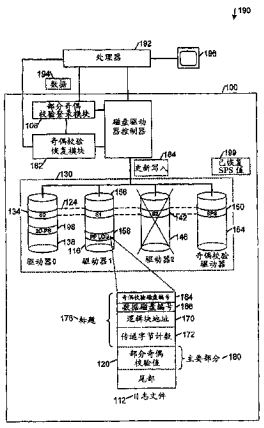
权利要求书 4 页 说明书 5 页 附图 3 页

(54)发明名称

电力安全磁盘存储装置、系统和方法

(57)摘要

装置和系统以及方法和制品可操作以在非易失性数据存储介质中创建日志文件。日志文件可包括标题记录、尾部记录以及与廉价磁盘冗余阵列(RAID)子系统中的条带关联的至少一个部分奇偶校验值。条带其中可包括位于可操作第一磁盘驱动器上的待更新的第一数据条、位于故障磁盘驱动器上的第二数据条以及位于可操作磁盘驱动器上的条带奇偶校验条。



CN 101512491 B

1. 一种恢复廉价盘冗余阵列 RAID 中的奇偶校验一致性的方法，包括：

在非易失性数据存储介质中创建日志文件，其中所述日志文件包括标题记录、尾部记录以及与廉价盘冗余阵列 RAID 子系统中的条带关联的至少一个部分奇偶校验值，并且其中所述条带包括位于可操作第一盘驱动器上的待更新的第一数据条、位于故障盘驱动器上的第二数据条以及位于可操作盘驱动器上的条带奇偶校验条，所述至少一个部分奇偶校验值是对第一数据条和所述条带奇偶校验条执行“异或”运算的结果；以及

执行恢复，包括对从所述日志文件获得的所述至少一个部分奇偶校验值和第一数据条执行“异或”运算。

2. 如权利要求 1 所述的方法，还包括：

从所述可操作第一盘驱动器中读取第一数据条；

从所述 RAID 子系统中读取所述条带奇偶校验条；

从第一数据条和所述条带奇偶校验条中计算所述至少一个部分奇偶校验值；

将奇偶校验盘编号、数据盘编号、逻辑块地址以及与第一数据条关联的传递字节计数写入所述日志文件中的所述标题记录；

将所述至少一个部分奇偶校验值写入所述日志文件的主要部分；以及

在将所述至少一个部分奇偶校验值写入所述主要部分之后，将所述标题记录复制到所述日志文件中的所述尾部记录。

3. 如权利要求 1 所述的方法，还包括：

在子系统意外复位之后比较所述标题记录与所述尾部记录，确定在发起对第一数据条和所述条带奇偶校验条的更新写入的时间与完成所述更新写入的时间之间的时段期间是否已经发生所述复位。

4. 如权利要求 3 所述的方法，还包括：

在所述子系统意外复位之后使用所述日志文件来恢复第一数据条与所述条带奇偶校验条之间的第一一致性。

5. 如权利要求 4 所述的方法，还包括：

从所述日志文件中读取所述至少一个部分奇偶校验值以及条带奇偶校验条和第一数据条位置值的集合；

从所述可操作第一盘驱动器中读取第一数据条；

对所述至少一个部分奇偶校验值和第一数据条执行第一奇偶校验恢复操作，以获得已恢复条带奇偶校验条值；以及

用所述已恢复条带奇偶校验条值来改写所述条带奇偶校验条。

6. 如权利要求 5 所述的方法，还包括：

如果在所述 RAID 子系统上存在使用第一数据条计算的第二维奇偶校验条，并且所述 RAID 子系统包括 RAID 级 6 子系统，则创建所述日志文件；

对所述已恢复条带奇偶校验条值和所述条带中的所有数据条执行第二奇偶校验恢复操作，以创建已恢复第二维奇偶校验条值；以及

用所述已恢复第二维奇偶校验条值来改写第二维奇偶校验条。

7. 如权利要求 1 所述的方法，其中所述条带奇偶校验条包括使用第一数据条和所述条带中包含的附加数据条计算的奇偶校验值。

8. 如权利要求1所述的方法，其中所述非易失性数据存储介质包括位于所述RAID子系统中包含的至少一个盘驱动器上的保留区。

9. 如权利要求1所述的方法，还包括：

聚集对所述日志文件的所述至少一个部分奇偶校验值的写入，以增加写入吞吐量。

10. 一种恢复廉价盘冗余阵列RAID中的奇偶校验一致性的产品，包括：

用于在非易失性数据存储介质中创建日志文件的单元，其中所述日志文件包括标题记录、尾部记录以及与廉价盘冗余阵列RAID子系统中的条带关联的至少一个部分奇偶校验值，并且其中所述条带包括位于可操作第一盘驱动器上的待更新的第一数据条、位于故障盘驱动器上的第二数据条以及位于可操作盘驱动器上的条带奇偶校验条，所述至少一个部分奇偶校验值是对第一数据条和所述条带奇偶校验条执行“异或”运算的结果；以及

用于执行恢复的单元，包括对从所述日志文件获得的所述至少一个部分奇偶校验值和第一数据条执行“异或”运算。

11. 如权利要求10所述的产品，还包括：

用于从所述可操作第一盘驱动器中读取第一数据条的单元；

用于从所述RAID子系统中读取所述条带奇偶校验条的单元；

用于从第一数据条和所述条带奇偶校验条中计算所述至少一个部分奇偶校验值的单元；

用于将奇偶校验盘编号、数据盘编号、逻辑块地址以及与第一数据条关联的传递字节计数写入所述日志文件中的标题记录的单元；

用于将所述至少一个部分奇偶校验值写入所述日志文件的主要部分的单元；以及

用于在将所述至少一个部分奇偶校验值写入所述主要部分之后将所述标题记录复制到所述日志文件中的尾部记录的单元。

12. 如权利要求11所述的产品，还包括：

用于将数字签名写入所述日志文件、以便用于确认来自所述日志文件的标题记录与所述至少一个部分奇偶校验值关联的单元。

13. 如权利要求10所述的产品，还包括：

用于在子系统意外复位之后比较所述标题记录与所述尾部记录、以确定在发起对第一数据条和所述条带奇偶校验条的更新写入的时间与完成所述更新写入的时间之间的时段期间是否已经发生所述复位的单元。

14. 如权利要求13所述的产品，还包括：

用于从所述日志文件中读取与所述至少一个部分奇偶校验值关联的数字签名的单元；以及

用于检验所述数字签名、以确认来自所述日志文件的所述标题记录与所述至少一个部分奇偶校验值关联的单元。

15. 如权利要求14所述的产品，还包括：

用于在子系统意外复位之后使用所述日志文件来恢复第一数据条与所述条带奇偶校验条之间的一致性的单元，其中在发起对第一数据条和所述条带奇偶校验条的所述更新写入的时间与完成所述更新写入的时间之间的时段期间发生所述复位。

16. 如权利要求15所述的产品，包括：

用于从所述日志文件中读取所述至少一个部分奇偶校验值以及条带奇偶校验条和第一数据条位置值的集合的单元；

用于从所述可操作第一盘驱动器中读取第一数据条的单元；

用于对所述至少一个部分奇偶校验值和第一数据条执行奇偶校验恢复操作、以获得已恢复条带奇偶校验条值的单元；以及

用于用所述已恢复条带奇偶校验条值来改写所述条带奇偶校验条的单元。

17. 一种恢复廉价盘冗余阵列 RAID 中的奇偶校验一致性的装置，包括：

部分奇偶校验登录模块，其在非易失性数据存储介质中创建日志文件，其中所述日志文件包括与廉价盘冗余阵列 RAID 子系统中的条带关联的至少一个部分奇偶校验值，并且其中所述条带包括位于可操作第一盘驱动器上的待更新的第一数据条、位于故障盘驱动器上的第二数据条以及位于可操作盘驱动器上的条带奇偶校验条，所述至少一个部分奇偶校验值是对第一数据条和所述条带奇偶校验条执行“异或”运算的结果；以及

奇偶校验恢复模块，耦合到所述登录模块，以通过对从所述日志文件获得的所述至少一个部分奇偶校验值和第一数据条执行“异或”运算来恢复第一数据条与所述条带奇偶校验条之间的一致性。

18. 如权利要求 17 所述的装置，其中所述日志文件还包括与所述条带奇偶校验条关联的奇偶校验盘编号、数据盘编号、逻辑块地址以及与第一数据条关联的传递字节计数。

19. 如权利要求 18 所述的装置，其中所述奇偶校验盘编号、所述数据盘编号、所述逻辑块地址以及所述传递字节计数位于所述日志文件的标题部分中，并且所述至少一个部分奇偶校验值位于所述日志文件的主要部分中。

20. 如权利要求 17 所述的装置，其中所述非易失性数据存储介质包括盘驱动器上的存储区。

21. 如权利要求 20 所述的装置，其中所述存储区包括位于所述 RAID 子系统中包含的至少一个盘驱动器上的保留区。

22. 如权利要求 17 所述的装置，所述奇偶校验恢复模块还包括：

以在子系统意外复位之后使用所述日志文件，其中在发起对第一数据条和所述条带奇偶校验条的更新写入的时间与完成所述更新写入的时间之间的时段期间发生所述复位，恢复所述一致性，所述奇偶校验恢复模块包括硬件模块、固件模块、或计算机可读介质上存储的软件模块中的至少一个。

23. 一种恢复廉价盘冗余阵列 RAID 中的奇偶校验一致性的系统，包括：

部分奇偶校验登录模块，其在非易失性数据存储介质中创建日志文件，其中所述日志文件包括与廉价盘冗余阵列 RAID 子系统中的条带关联的至少一个部分奇偶校验值，并且其中所述条带包括位于可操作第一盘驱动器上的待更新的第一数据条、位于故障盘驱动器上的第二数据条以及位于可操作盘驱动器上的条带奇偶校验条，所述至少一个部分奇偶校验值是对第一数据条和所述条带奇偶校验条执行“异或”运算的结果；以及

奇偶校验恢复模块，耦合到所述登录模块，以在子系统意外复位之后使用所述日志文件来恢复第一数据条与所述条带奇偶校验条之间的第一一致性，其中在发起对第一数据条和所述条带奇偶校验条的更新写入的时间与完成所述更新写入的时间之间的时段期间发生所述复位，所述奇偶校验恢复模块通过对从所述日志文件获得的所述至少一个部分奇偶

校验值和第一数据条执行“异或”运算来恢复第一一致性。

24. 如权利要求 23 所述的系统，其中所述 RAID 子系统包括 RAID 级 3、RAID 级 4、RAID 级 5、RAID 级 6、RAID 级 30 和 RAID 级 50 中的至少一个。

25. 如权利要求 23 所述的系统，还包括：

第二维奇偶校验条，位于可操作盘驱动器上并使用第一数据条进行计算，其中所述 RAID 子系统包括 RAID 级 6 子系统，并且其中所述奇偶校验恢复模块使用已恢复条带奇偶校验条和所述条带中的所有数据条来恢复第一数据条与第二维奇偶校验条之间的第二一致性。

26. 如权利要求 23 所述的系统，其中所述非易失性数据存储介质包括位于所述 RAID 子系统中包含的至少一个盘驱动器上的保留区。

27. 如权利要求 23 所述的系统，其中所述子系统意外复位包括：RAID 子系统处理的中断，而在所述中断之前没有通知所述 RAID 子系统。

电力安全磁盘存储装置、系统和方法

技术领域

[0001] 本文所述的各种实施例一般涉及磁盘存储技术，包括用于在电力丢失之后恢复冗余阵列中的奇偶校验一致性的装置、系统和方法。

背景技术

[0002] 磁盘存储系统可包括廉价磁盘冗余阵列 (RAID) 子系统。RAID 子系统可存储位于 RAID 子系统中包含的若干数据磁盘驱动器的每个上称作“条 (strip)”的数据聚集。来自每个数据磁盘驱动器的各条可通过位于若干数据磁盘驱动器之一上或者位于单独奇偶校验磁盘驱动器上的奇偶校验条与其它条相关。奇偶校验条和相关数据条可统称为“条带 (stripe)”。

[0003] 奇偶校验条可通过对相关数据条执行“异或”运算来进行计算，并且可与数据条一起由 RAID 子系统用来重新计算在写入条带之后的某个时间变为不可读的数据条的值。因此，可将冗余结合到 RAID 子系统中，并且冗余可实现在一个或多个磁盘驱动器故障之后恢复卷（操作的“降级模式”）。但是，如果对目标数据条或者奇偶校验条执行写入更新的处理器或控制器正以降级模式进行操作时意外地复位了，则可能出现问题。意外复位可能例如由电力丢失或操作系统致命错误而触发。奇偶校验条在复位之后可能与条带中的数据条不一致。

附图说明

[0004] 图 1 是根据本发明各种实施例的装置和典型系统的框图。

[0005] 图 2 是说明根据本发明各种实施例的若干方法的流程图。

[0006] 图 3 是根据本发明各种实施例的制品的框图。

具体实施方式

[0007] 图 1 包括根据本发明各种实施例的装置 100 和系统 190 的框图。装置 100 可包括在非易失性数据存储介质 116 中创建日志文件 112 的部分奇偶校验登录模块 106。日志文件 112 可包括与 RAID 子系统 130 中的条带 124 关联的一个或多个部分奇偶校验值 120。RAID 子系统 130 可包括 RAID 级 3、RAID 级 4、RAID 级 5、RAID 级 6、RAID 级 30、RAID 级 50 或采用成条数据和从组成条带的数据条集合中计算的一个或多个奇偶校验值的任何其它 RAID 或冗余磁盘阵列子系统。

[0008] 条带 124 其中可包括位于可操作第一磁盘驱动器 138 上的待更新的第一数据条 134、位于故障磁盘驱动器 146 上的第二数据条 142 以及位于可操作磁盘驱动器 154 上的条带奇偶校验条 150。部分奇偶校验值 120 可包括对第一数据条 134 和条带奇偶校验条 150 执行“异或”运算的结果。非易失性数据存储介质 116 可包括磁盘驱动器 156 上的存储区，其中包括位于 RAID 子系统 130 所包含的一个或多个磁盘驱动器上的保留区 158。

[0009] 日志文件 112 还可包括与条带奇偶校验条 150 关联的奇偶校验磁盘编号 164、数据

磁盘编号 166、逻辑块地址 170 以及与第一数据条 134 关联的传递字节计数 172。后面的奇偶校验和数据条位置信息可存在于日志文件的标题部分 176；且部分奇偶校验值 120 可存在于日志文件的主要部分 180。这些值可标识使用来自日志文件的部分奇偶校验值参与奇偶校验恢复操作的磁盘驱动器。

[0010] 装置 100 还可包括耦合到登录模块 106 的奇偶校验恢复模块 182，以便恢复第一数据条 134 与条带奇偶校验条 150 之间的一致性。在子系统意外复位（例如，在发起对第一数据条 134 和条带奇偶校验条 150 的更新写入 184 的时间与完成更新写入 184 的时间之间发生复位）之后，可能需要恢复一致性。恢复可包括对从日志文件 112 得到的部分奇偶校验值 120 和第一数据条 134 执行“异或”运算。

[0011] 可实现其它实施例。例如，系统 190 可包括一个或多个装置 100，其中包括部分奇偶校验登录模块 106、日志文件 112 和非易失性数据存储介质 116。非易失性数据存储介质 116 可包括位于 RAID 子系统 130 所包含的一个或多个磁盘驱动器上的保留区 158。系统 190 还可包括如对于装置 100 所述的一个或多个部分奇偶校验值 120、条带 124、第一数据条 134、条带奇偶校验条 150、奇偶校验恢复模块 182 以及 RAID 子系统 130。RAID 子系统 130 可包括 RAID 级 3、RAID 级 4、RAID 级 5、RAID 级 6、RAID 级 30、RAID 级 50 或者采用成条数据和从组成条带的数据条集合中计算的一个或多个奇偶校验值的任何其它 RAID 或冗余磁盘阵列子系统。

[0012] 系统 190 还可包括：一个或多个处理器 192，耦合到登录模块 106，提供写入第一数据条 134 的数据 194；以及显示器 196，耦合到处理器 192，显示数据 194。显示器 196 可包括阴极射线管显示器或者固态显示器，诸如液晶显示器、等离子显示器、或者发光二极管显示器等等。

[0013] 系统 190 的一些实施例可包括位于 RAID 级 6 子系统内的可操作磁盘驱动器 138 上的第二维奇偶校验条 198。可使用第一数据条 134 以及条带 124 中的其它数据条、可能包括所有其它数据条，来计算第二维奇偶校验条 198。奇偶校验恢复模块 182 可用于在子系统意外复位（例如，RAID 子系统处理中断而在中断之前没有通知 RAID 子系统）之后恢复第一数据条 134 与第二维奇偶校验条 198 之间的第二一致性。第二维奇偶校验一致性恢复操作可包括对已恢复条带奇偶校验条值 199 和条带 124 中的其它数据条执行的操作。

[0014] 前面所述的任一个组件可通过许多方式来实现，包括通过软件的实施例。因此，装置 100、部分奇偶校验登录模块 106、日志文件 112、非易失性数据存储介质 116、部分奇偶校验值 120、条带 124、廉价磁盘冗余阵列 (RAID) 子系统 130、数据条 134、142、磁盘驱动器 138、146、154、156、条带奇偶校验条 150、保留区 158、磁盘编号 164、166、逻辑块地址 170、传递字节计数 172、日志文件的部分 176、180、奇偶校验恢复模块 182、更新写入 184、系统 190、处理器 192、数据 194、液晶显示器 196、第二维奇偶校验条 198 以及已恢复条带奇偶校验条值 199 在本文中均可表示为“模块”。

[0015] 根据装置 100 和系统 190 的设计者的需要以及适合于各种实施例的具体实现，这些模块可包括硬件电路、单或多处理器电路、存储器电路、软件程序模块和对象、固件以及它们的组合。因此，模块可包含在诸如软件电信号模拟包、功率使用和分配模拟包、电容电感模拟包、功率 / 热耗散模拟包、信号发射接收模拟包之类的系统操作模拟包中，或者用来模拟各种可能实施例操作的软件和硬件的任何组合中。例如，这些模拟可用于表征或测试

实施例。

[0016] 还应当理解，各种实施例的装置和系统可用于与在意外复位之后恢复降级模式RAID卷的条带奇偶校验一致性不同的应用。因此，本发明的各种实施例不受此限制。装置100和系统190的说明用于提供对各种实施例结构的总体理解，并且它们不是意在用作对于可能利用本文所述结构的装置和系统的全部要素及特征的完整描述。

[0017] 可包括各种实施例的新颖装置和系统的应用包括用于高速计算机的电子电路、通信和信号处理电路、调制解调器、单或多处理器模块、单或多嵌入式处理器、数据开关以及包含多层多芯片模块的专用模块。这类装置和系统还可作为子组件包含在诸如电视、蜂窝电话、个人计算机、工作站、无线电、视频播放器、车辆等各种电子系统中。一些实施例可包括多种方法。

[0018] 图2是说明根据本发明各种实施例的若干方法211的流程图。一种这样的方法211可在框223以在非易失性数据存储介质中创建日志文件开始。非易失性数据存储介质可包括位于RAID子系统所包含的一个或多个磁盘驱动器上的保留区。如前面所述的，RAID子系统可包括RAID级3、RAID级4、RAID级5、RAID级6、RAID级30、RAID级50或者采用成条数据和从组成条带的数据条集合中计算的一个或多个奇偶校验值的任何其它RAID或冗余磁盘阵列子系统。

[0019] 日志文件可包括标题记录、尾部记录以及与RAID子系统中的条带关联的一个或多个部分奇偶校验值。条带其中可包括位于可操作第一磁盘驱动器上的待更新的第一数据条、位于故障磁盘驱动器上的第二数据条以及位于可操作磁盘驱动器上的条带奇偶校验条。条带奇偶校验条可包括使用第一数据条和条带中包含的附加数据条所计算的奇偶校验值。

[0020] 方法211可继续进行在框231从可操作第一磁盘驱动器中读取第一数据条，以及在框237从RAID子系统中读取条带奇偶校验条。方法211还可包括在框241可能通过对第一数据条和条带奇偶校验条执行“异或”功能，从第一数据条和条带奇偶校验条中计算一个或多个部分奇偶校验值。方法211还可包括在框245将奇偶校验磁盘编号、数据磁盘编号、逻辑块地址以及与第一数据条关联的传递字节计数写入日志文件的标题记录。

[0021] 方法211可继续进行在框251聚集对日志文件的部分奇偶校验值的写入以增加写入吞吐量，以及在框255将部分奇偶校验值写入日志文件的主要部分。方法211还可包括在框259在将部分奇偶校验值写入主要部分之后将标题记录复制到日志文件的尾部记录，以及在框261将数字签名写入日志文件，以用于确认来自日志文件的标题记录与部分奇偶校验值关联。

[0022] 对数据条集合的更新集合可在完成与更新集合对应的对日志文件的写入时被写入磁盘，其中包括将日志文件标题内容复制到日志文件尾部部分。这样，标题和尾部的比较可提供关于数据条磁盘写入集合正在进行还是尚未开始的指示。数据条磁盘写入集合可被认为在与数据条更新集合对应的对日志文件的部分奇偶校验写入完成之后的某个时间已经开始，并且标题内容被复制到尾部部分，从而使标题和尾部相等。因此，如果标题和尾部不相等，则日志文件可被认为正进行更新，并且数据条写入集合可能还没开始。

[0023] 方法211还可继续进行在框263在子系统意外复位之后比较标题记录与尾部记录，以便确定在发起对第一数据条和条带奇偶校验条的更新写入的时间与完成更新写入的

时间之间是否已经发生复位。如果奇偶校验一致性需要恢复，则方法 211 可继续进行在框 265 从日志文件中读取与部分奇偶校验值关联的数字签名，以及在框 266 检验数字签名以便确认来自日志文件的标题记录与部分奇偶校验值关联。

[0024] 方法 211 还可包括在框 267 在子系统意外复位之后使用日志文件来恢复第一数据条与条带奇偶校验条之间的第一一致性。恢复第一一致性可继续进行在框 271 从日志文件中读取部分奇偶校验值以及条带奇偶校验条和第一数据条位置值的集合。方法 211 可继续进行在框 275 从可操作第一磁盘驱动器中读取第一数据条。方法 211 可包括在框 279 对部分奇偶校验值和第一数据条执行可能包括“异或”运算的第一奇偶校验恢复操作，以获得已恢复条带奇偶校验条值。方法 211 还可包括在框 283 用已恢复条带奇偶校验条值来改写条带奇偶校验条。

[0025] 日志文件可在 RAID 级 6 子系统中创建，其中使用第一数据条来计算第二维奇偶校验条。方法 211 可包括在框 287 对已恢复条带奇偶校验条值和条带中的所有数据条执行第二奇偶校验恢复操作，以便创建已恢复第二维奇偶校验条值。方法 211 可在框 291 以用已恢复第二维奇偶校验条值改写第二维奇偶校验条来结束。

[0026] 本文所述的方法不必按照所述的顺序或者任何特定顺序来执行。此外，针对本文确定的方法所述的各种活动可以重复、串行或并行方式执行。包括参数、命令、操作数及其它数据的信息可采取一个或多个载波的形式来发送和接收。

[0027] 本领域的技术人员会理解可从基于计算机的系统中的计算机可读介质中运行软件程序以便执行软件程序中定义的功能的方式。各种编程语言可用来创建设计成实现和执行本文公开的方法的一个或多个软件程序。程序可以面向对象的格式使用面向对象的语言、如 Java 或 C++ 来构造。或者，程序可以面向过程的格式使用过程语言、如汇编或 C 来构造。软件组件可使用本领域技术人员众所周知的多种机制进行通信，例如应用程序接口或进程间通信技术，包括远程过程调用。各种实施例的示教不限于任何特定的编程语言或环境。因此，可实现其它实施例，如下面针对图 3 所述的。

[0028] 图 3 是根据本发明各种实施例的制品 385 的框图。这类实施例的实例可包括计算机、存储系统、磁盘或光盘、某种其它存储装置或者任何类型的电子设备或系统。制品 385 可包括耦合到机器可存取介质、如存储器 389（例如包括电、光或电磁元件的存储器）的一个或多个处理器 387。介质可包含在被存取时使机器（例如处理器 387）在非易失性数据存储介质中创建日志文件的关联信息 391（例如计算机程序指令、数据或者它们两者），其中日志文件包括标题记录、尾部记录以及与廉价磁盘冗余阵列 (RAID) 子系统中的条带关联的一个或多个部分奇偶校验值。条带可包括位于可操作第一磁盘驱动器上的待更新的第一数据条、位于故障磁盘驱动器上的第二数据条以及位于可操作磁盘驱动器上的条带奇偶校验条。

[0029] 其它活动可包括从可操作第一磁盘驱动器中读取第一数据条、从 RAID 子系统中读取条带奇偶校验条、以及从第一数据条和条带奇偶校验条中计算部分奇偶校验值。活动还可包括将奇偶校验磁盘编号、数据磁盘编号、逻辑块地址以及与第一数据条关联的传递字节计数写入日志文件的标题记录，以及将部分奇偶校验值写入日志文件的主要部分。其它活动可包括在将部分奇偶校验值写入主要部分之后将标题记录复制到日志文件的尾部记录，以及将数字签名写入日志文件以便用于确认来自日志文件的标题记录与部分奇偶校

验值关联。

[0030] 活动还可包括在子系统意外复位之后比较标题记录与尾部记录，以便确定在发起对第一数据条和条带奇偶校验条的更新写入的时间与完成更新写入的时间之间的时段期间是否已经发生复位。附加活动可包括从日志文件中读取与部分奇偶校验值关联的数字签名，并检验该数字签名以确认来自日志文件的标题记录与部分奇偶校验值关联。

[0031] 其它活动可包括在子系统意外复位之后使用日志文件来恢复第一数据条与条带奇偶校验条之间的一致性。这些活动可包括从日志文件中读取部分奇偶校验值以及条带奇偶校验条和第一数据条位置值的集合。活动还可包括：从可操作第一磁盘驱动器中读取第一数据条；对部分奇偶校验值和第一数据条执行奇偶校验恢复操作，以便获得已恢复条带奇偶校验条值；以及用已恢复条带奇偶校验条值来改写条带奇偶校验条。

[0032] 实现本文公开的装置、系统和方法可操作以在对与条带关联的数据或奇偶校验条的更新期间的意外复位之后，恢复降级模式 RAID 卷中的条带奇偶校验一致性。

[0033] 构成本主题的组成部分的附图通过说明而非限制来说明可实施本主题的具体实施例。充分详细地描述了所示实施例，使本领域的技术人员能够实施本文公开的示教。可以使用并可从中导出其它实施例，使得可进行结构和逻辑替换及改变，并不背离本公开的范围。因此，该具体实施方式不应认为是限制性的，并且各种实施例的范围仅由所附权利要求书以及这种权利要求书涵盖的全部等效范围来限定。

[0034] 发明主题的这类实施例在本文中可单独或共同由术语“本发明”来表示，只是为了方便，而不是要有意将本申请的范围限制于任何单一发明或发明概念，如果实际上公开了不止一个的话。因此，虽然本文已经说明和描述了具体实施例，但是，用于实现相同目的的任何配置均可替代所示的具体实施例。本公开意在涵盖各种实施例的任何以及全部修改和变更。通过阅读以上说明，本领域的技术人员会非常清楚以上实施例以及本文没有明确描述的其它实施例的组合。

[0035] 公开的摘要是根据 37 C.F.R. §1.72(b) 要求摘要以使读者快速了解技术公开的本质而提供的。应当理解，它的提供并不是用于解释或限制权利要求书的范围或含意。另外，在以上具体实施方式中可以看到，各种特征集中在单个实施例中，用于简化本公开。这个公开的方法不应解释为要求超过各权利要求中明确描述的特征。而是，发明主题可存在于少于单个公开实施例的全部特征。因此，以下权利要求书由此结合到具体实施方式中，其中各权利要求本身代表单独的实施例。

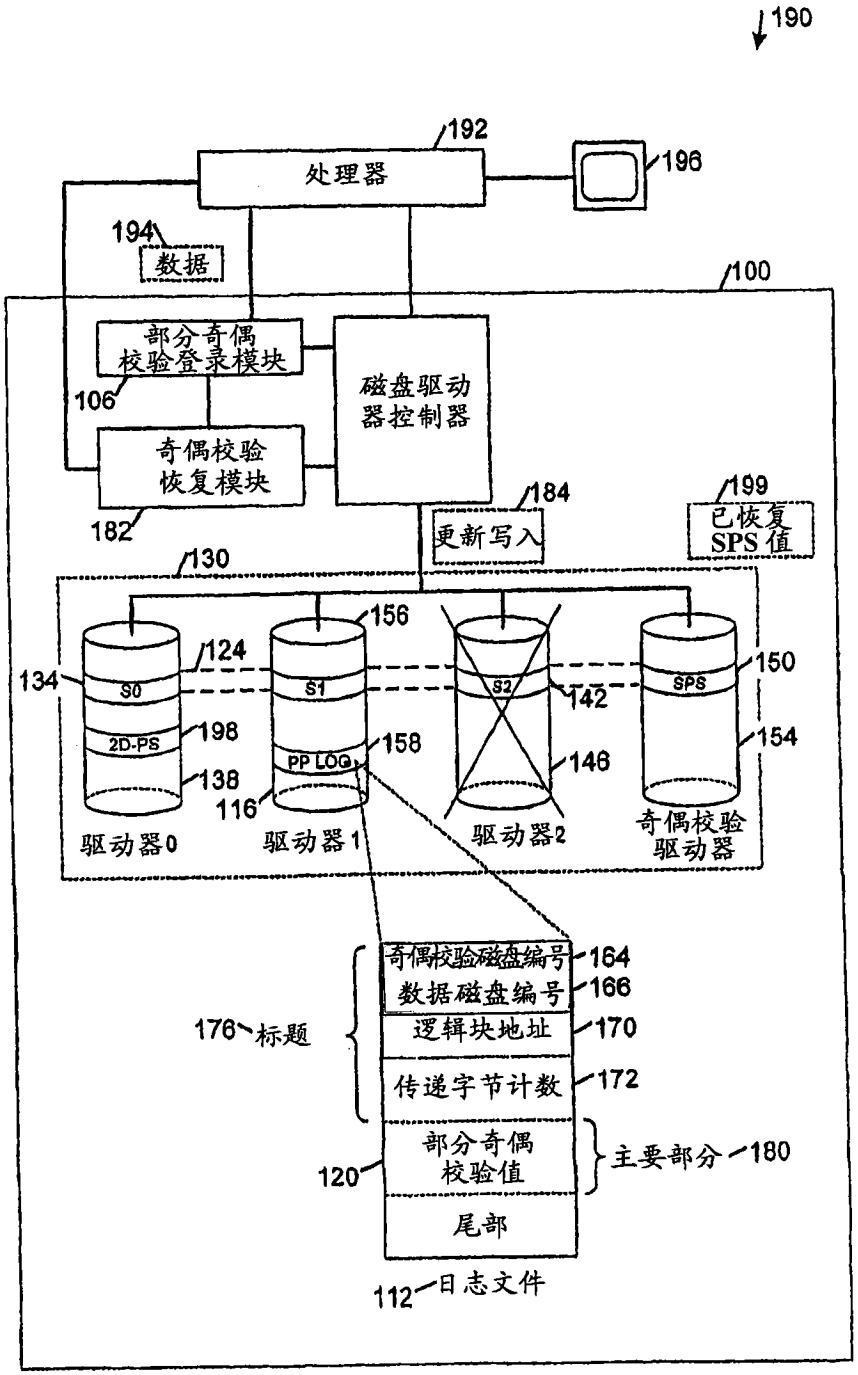
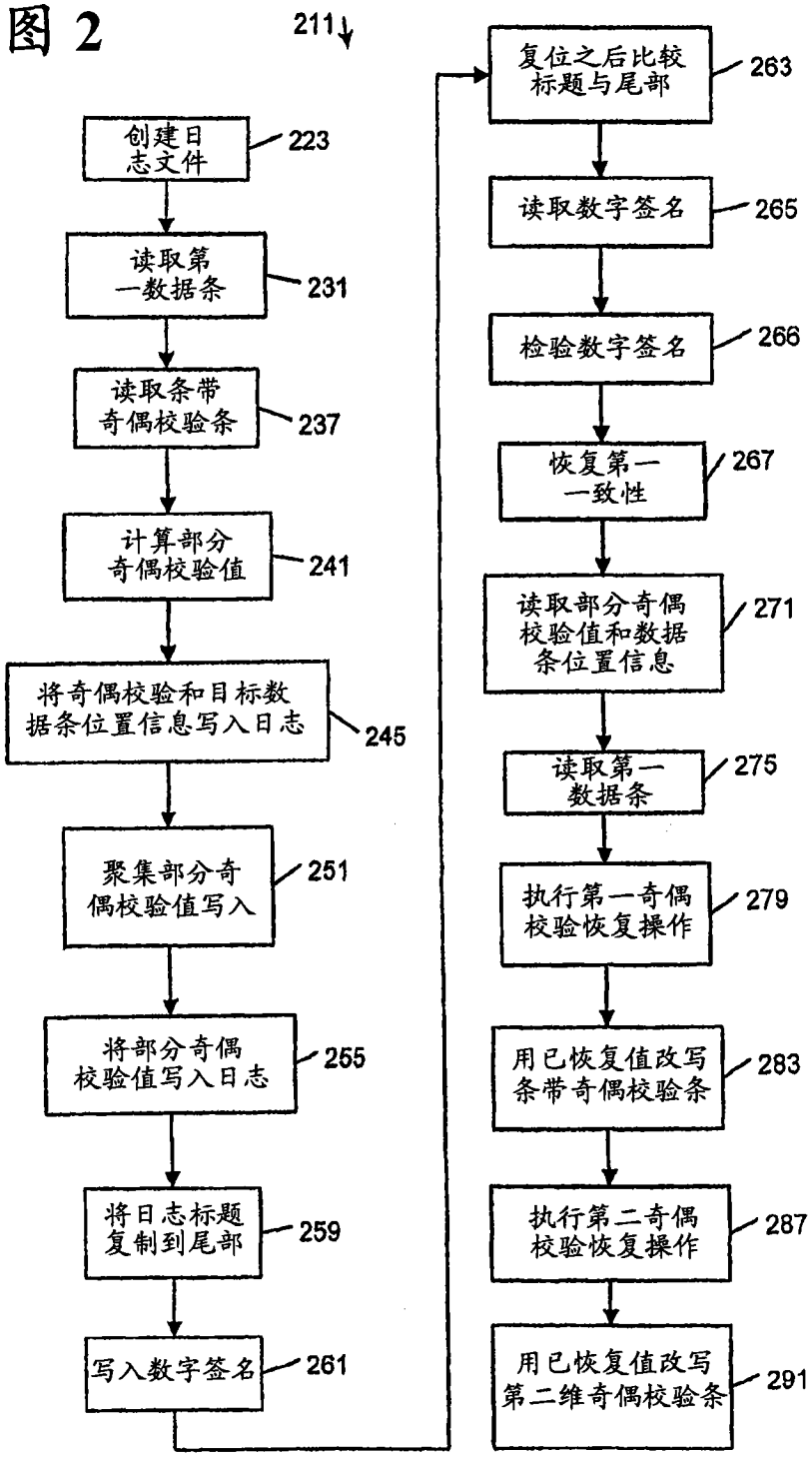

190
192
处理器
196
194
数据
100
部分奇偶
校验登录模块
106
磁盘驱动
器控制器
奇偶校验
恢复模块
182
184
更新写入
199
已恢复
SPS 值
130
156
124
134
S0
S1
S2
SPS
150
142
2D-PS
198
PP LOG
158
146
154
138
116
驱动器0
驱动器1
驱动器2
奇偶校验
驱动器
奇偶校验磁盘编号
164
数据磁盘编号
166
176
标题
逻辑块地址
170
传递字节计数
172
部分奇偶
校验值
主要部分
180
120
尾部
112
日志文件

图 1

图 2


211
创建日志文件
223
读取第一数据条
231
读取条带奇偶校验条
237
计算部分奇偶校验值
241
将奇偶校验和目标数据条位置信息写入日志
245
聚集部分奇偶校验值写入
251
将部分奇偶校验值写入日志
255
将日志标题复制到尾部
259
写入数字签名
261
复位之后比较标题与尾部
263
读取数字签名
265
检验数字签名
266
恢复第一一致性
267
读取部分奇偶校验值和数据条位置信息
271
读取第一数据条
275
执行第一奇偶校验恢复操作
279
用已恢复值改写条带奇偶校验条
283
执行第二奇偶校验恢复操作
287
用已恢复值改写第二维奇偶校验条
291

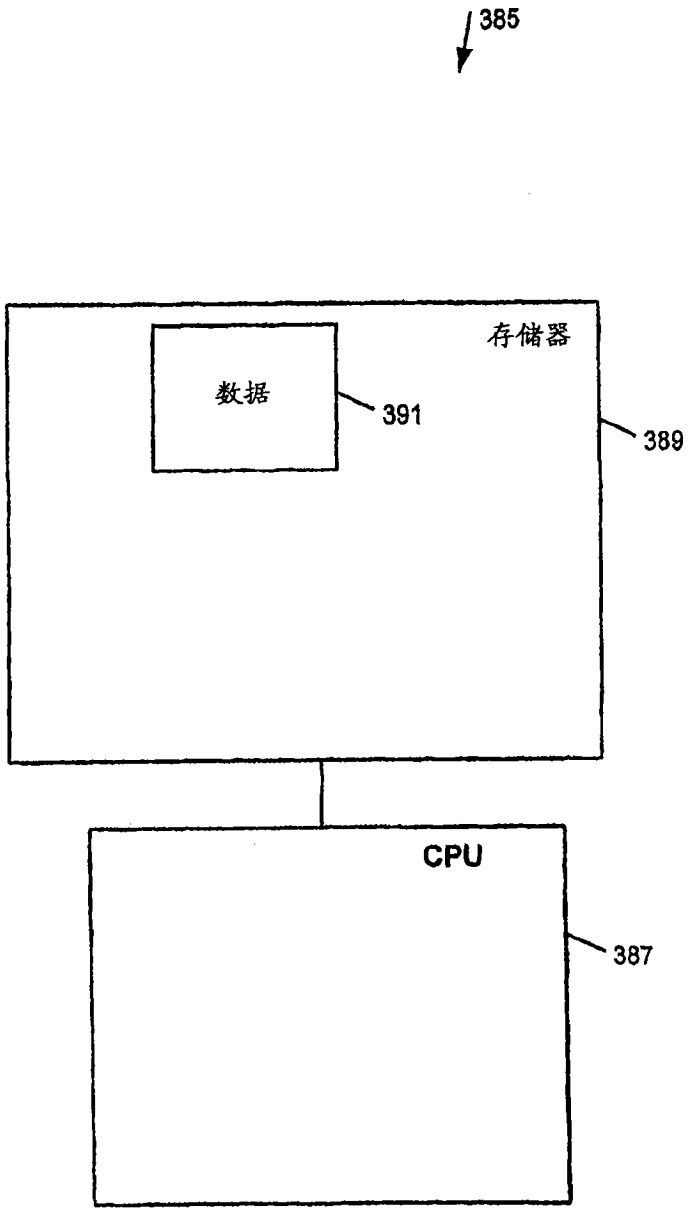


图 3