



(12) 发明专利申请

(10) 申请公布号 CN 105074654 A

(43) 申请公布日 2015. 11. 18

(21) 申请号 201480019299. 3

代理人 赵蓉民 徐东升

(22) 申请日 2014. 03. 31

(51) Int. Cl.

(30) 优先权数据

G06F 9/30(2006.01)

61/807, 673 2013. 04. 02 US

14/061, 592 2013. 10. 23 US

(85) PCT国际申请进入国家阶段日

2015. 09. 29

(86) PCT国际申请的申请数据

PCT/US2014/032418 2014. 03. 31

(87) PCT国际申请的公布数据

W02014/165454 EN 2014. 10. 09

(71) 申请人 西部数据技术公司

地址 美国加利福尼亚州

(72) 发明人 N·萨德里

(74) 专利代理机构 北京纪凯知识产权代理有限

公司 11245

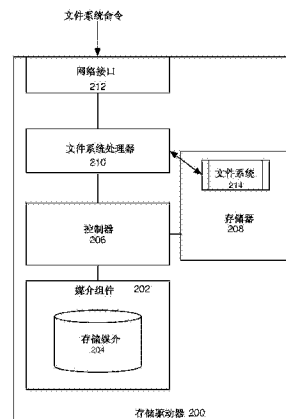
权利要求书2页 说明书5页 附图4页

(54) 发明名称

用于存储设备上的文件系统命令的特权化执行支持的方法和系统

(57) 摘要

本发明涉及一种存储设备,其能够通过对各种资源(诸如,存储媒介、硬件、存储器、固件等)的特权化存取执行更高水平的命令,诸如网络级文件系统命令。在一个实施例中,存储设备经配置以接收和执行网络级文件系统命令,诸如服务器消息块协议命令。具体地,存储设备包括具有存储媒介和通信接口(诸如网络接口和控制器)的驱动器。控制器经配置以解释和执行在存储媒介上存储的数据上的从通信接口接收的网络级文件系统命令。因此,存储设备可以更有效地服务网络级文件系统命令,而无需用户空间应用。



1. 一种经配置执行文件系统命令的存储驱动器,所述存储驱动器包括:  
至少一个存储媒介;  
控制器,其经配置以控制所述至少一个存储媒介;  
存储器,其耦合至所述控制器用于存储在所述至少一个存储媒介上存储的文件的至少一部分文件系统;  
通信接口,其经配置以与网络通信并接收网络级文件系统命令;以及  
处理器,其耦合到所述通信接口并且具有对所述控制器和至少一个存储媒介的特权化存取,并且由程序代码配置从而执行涉及在所述至少一个存储媒介上存储的文件的所述接收的网络级文件系统命令。
2. 根据权利要求1所述的存储驱动器,其中所述至少一个存储媒介包括硬盘。
3. 根据权利要求1所述的存储驱动器,其中所述至少一个存储媒介包括固态介质。
4. 根据权利要求1所述的存储驱动器,其中所述通信接口包括以太网协议网络通信接口。
5. 根据权利要求4所述的存储驱动器,其中所述处理器执行网络级文件系统命令。
6. 根据权利要求4所述的存储驱动器,其中所述处理器执行服务器消息块命令。
7. 根据权利要求1所述的存储驱动器,其中所述控制器由程序代码配置用于嵌入式操作系统。
8. 根据权利要求1所述的存储驱动器,其中所述处理器包括实施程序代码从而执行所述网络级文件系统命令的固件。
9. 一种存储设备,其经配置以通过对所述存储设备的存储媒介进行特权化存取从而执行命令,所述存储设备包括:  
至少一个存储媒介;  
控制器,其经配置以控制所述至少一个存储媒介;  
存储器,其耦合至所述控制器用于存储涉及在所述至少一个存储媒介上存储的文件的至少一部分文件系统;  
网络接口,其经配置以为网络提供接口;以及  
处理器,其执行操作系统,具有对所述控制器和所述至少一个存储媒介的特权化存取,并被耦合到所述网络接口,并且由程序代码配置从而执行涉及在所述至少一个存储媒介上存储的文件的网络级文件系统命令。
10. 根据权利要求9所述的存储设备,其中所述存储媒介包括至少一个硬盘。
11. 根据权利要求9所述的存储设备,其中所述控制器经配置以执行嵌入式Linux操作系统。
12. 根据权利要求9所述的存储设备,其中所述控制器经配置以执行嵌入式Windows操作系统。
13. 根据权利要求9所述的存储设备,其中所述处理器包括实施所述程序代码的固件从而执行网络级文件系统命令。
14. 根据权利要求9所述的存储设备,其中所述处理器经配置以基于对所述至少一个存储媒介的特权化存取,执行在所述至少一个存储媒介上存储的文件上的服务器消息块协议命令。

15. 一种通过存储驱动器处理网络级文件系统命令的方法,其中所述存储驱动器包括存储媒介、用于控制所述存储媒介的控制器、耦合到所述控制器且存储涉及在存储媒介上存储的文件的至少一部分文件系统的存储器、用于与网络通信的网络通信接口、以及基于对所述控制器和所述存储驱动器的存储媒介的特权化存取执行程序代码的处理器,所述方法包括:

经由所述网络通信接口接收至少一个网络级文件系统命令;

将所述接收的至少一个网络级文件系统命令从所述网络接口转发到所述处理器;以及

基于所述处理器对所述控制器和所述存储媒介的所述特权化存取,通过所述存储媒介上的所述处理器执行所述网络级文件系统命令。

16. 根据权利要求 15 所述的方法,其中执行所述网络级文件系统命令包括在所述处理器内的固件内执行所述网络级文件系统命令。

17. 根据权利要求 15 所述的方法,其中执行所述网络级文件系统命令包括通过所述处理器执行所述存储媒介上的服务器消息块协议命令。

18. 一种经配置以执行网络级文件系统命令的网络连接的存储设备,所述存储设备包括:

存储媒介;

控制器,其用于控制所述存储媒介;

存储器,其存储在该存储媒介上存储的文件的至少一部分文件系统;

网络接口,其经配置以为网络提供接口并接收网络级文件系统命令;以及

处理器,其耦合到所述网络接口,通过对所述控制器和存储媒介的特权化存取执行程序代码,从而执行涉及在该存储媒介上存储的文件的网络级文件系统命令。

19. 根据权利要求 18 所述的存储设备,其中所述存储媒介包括至少一种硬盘驱动器。

20. 根据权利要求 18 所述的存储设备,其中所述控制器经配置以执行嵌入式 LINUX 操作系统。

21. 根据权利要求 18 所述的存储设备,其中所述处理器经配置以执行 Windows 操作系统。

## 用于存储设备上的文件系统命令的特权化执行支持的方法和系统

### 背景技术

[0001] 数据存储实际上是所有计算系统的基本功能。目前,存在有大量可用的数据存储设备和技术。然而,通常所有数据存储涉及数据存储设备或驱动器(诸如硬盘驱动器、固态驱动器等)的使用。已知类型的驱动器在块级(block level)上存储并检索信息并且使用某些形式的块级存取(诸如 SCSI、IDE/ATA、SAS、SATA 等)为已知类型的驱动器提供接口。使用电缆通常可以直接将驱动器附接到客户端或主计算设备,并且因此所述驱动器通常被称为直接外接存储或“DAS”。为了进一步扩张存储的可用性,可以经由网络诸如使用 iSCSI、网络外接存储(“NAS”)或存储区网络(“SAN”)来存取一个或多个驱动器。

[0002] 然而,信息通常是基于文件系统来组织的。文件系统,诸如网络文件系统(“NFS”)、通用英特网文件系统(“CIFS”)、服务器消息块(“SMB”)等都是已知的文件系统,上述已知的文件系统采用网络级文件系统命令,经由网络(诸如基于英特网协议(“IP”)的网络)提供对文件等的共享存取。不幸的是网络级文件系统命令不是由存储驱动器本地处理的。

[0003] 而是,存储驱动器必须与执行并翻译网络级文件系统命令为块级命令的另一个设备或计算资源耦合。DAS 驱动器依靠在与操作系统和应用级软件相对应的主机(诸如个人计算机、服务器等)上运行的操作系统和应用级软件。NAS 设备包括执行操作系统(诸如 LINUX 或 Windows)的控制器,所述操作系统支持运行可以处理网络级文件系统命令的应用。iSCSI 和 SAN 存储系统依靠应用服务器以处理文件系统命令。

[0004] 例如,文件共享软件, SAMBA 是公知的应用,其可以接收文件系统命令并且为存储驱动器提供接口。然而,该软件在用户空间中运行,并使用设备驱动器以及由在存储设备上运行的操作系统提供的其它服务。不幸的是,用户空间软件驱动器可能不稳定和/或经受低性能的问题。该性能缺点在嵌入式计算设备和存储设备上特别显著。因此,需要提供改善计算设备特别是存储系统性能的方法和系统。

### 附图说明

[0005] 下面参考附图描述体现本发明各种特征的系统和方法,其中:

[0006] 图 1A 和 1B 示出采用用户空间应用支持文件系统命令的常规存储设备。

[0007] 图 2 示出示例性存储设备,其支持对存储设备的资源进行特权化的存取从而执行文件系统命令。

[0008] 图 3 示出根据本发明实施例的示例性过程。

### 具体实施方式

[0009] 本实施例涉及优化对存储文件的存取,特别是在网络环境中对存储文件的存取。在一个实施例中,存储驱动器能够通过通过对各种资源(诸如,存储媒介、硬件、存储器、固件等)进行特权化存取从而执行更高水平的命令,诸如网络级的文件系统命令。在一个实施例中,存储驱动器经配置以接收和执行网络级文件系统命令,诸如 NFS 或 SMB 命令。因此,

存储驱动器可以更有效地直接接收和执行网络级的文件系统命令，而无需执行用户空间应用。

[0010] 本发明的实施例可以在多种操作系统（诸如 LINUX、Windows、UNIX、Mac OSX 等）中的任何一种上实施。为了以说明为目的，示出了运行 LINUX 操作系统的存储设备的示例性实施例。一个实施例可以采用特殊目的硬件或固件组件以执行文件系统命令。可替换地，另一个实施例使用内核模式设备驱动器以支持网络文件系统命令。

[0011] 此外，实施例可以通过存储媒介而被实施在任何类型存储驱动器上，该存储媒介包括硬盘媒介、固态媒介、闪存等。在实施例中，网络级文件系统命令是传输不可知的，且因此可以在以太网、Wi-Fi、无线带宽（infiniband）等上传输。网络级文件系统协议可以包括但不限于 NFS、SMB、苹果文件协议（“AFP”）、文件传输协议（“FTP”）、远程文件系统（“RFS”）、WebDAV 等。

[0012] 下面将描述本发明的某些实施例。这些实施例是仅通过示例的方式呈现的，而不旨在限制本发明的范围。实际上，这里所说的新颖方法和系统可以通过各种其它的形式来体现。而且，在不偏离本发明精神的情况下，对在此所述的方法和系统的形式做出不同省略、替换和变化。为了示出某些实施例，将对附图做出参考。

[0013] 图 1A 和 1B 显示常规存储设备，其提供经由网络的网络共享。具体地，图 1A 示出常规直接附接存储驱动器。图 1B 示出常规网络附接存储。

[0014] 现在参考图 1A，存储驱动器 102 可以包括媒介组件 104，所述媒介组件 104 具有存储媒介 106、驱动器控制器 108、存储器 110、以及驱动器接口 112。这些组件是本领域技术人员公知的。

[0015] 如图所示，存储设备 102 被连接到客户端设备 114。作为常规驱动器，存储设备 102 是块级设备并经由存储接口 116 为客户端设备 114 提供接口。例如，存储接口 116 和驱动器接口 112 可以是 SCSI、IDE/ATA、SAS、或 SATA 接口。

[0016] 为了实施文件系统，客户端设备 114 还包括支持执行一个或多个应用 120 的操作系统 118。在操作过程中，由操作系统 118 控制且维持的应用 120 基于文件和文件系统来操作。为了存取或检索文件，操作系统 118 识别块与哪一个特定文件相对应并经由驱动器接口 112 将各种块级命令提供给存储驱动器 102。随后，控制器 108 在存储媒介 106 上定位请求块并指示媒介组件 104。因此，在该基本存储配置中，在应用级通过运行在操控系统的用户空间中的应用来操控文件系统命令。

[0017] 然而，相同类型的方法经由网络被用在共享文件系统或分布式文件系统中。如图 1B 所示，其显示网络附接的存储（NAS）100。NAS 100 可以包括一个或多个存储驱动器 102，例如，在底座或机架安装机箱中。此外，NAS 100 包括 NAS 控制器 116，其执行操作系统 118 和文件服务器应用 120。NAS 100 进一步包括网络接口 120 以连接网络 122。

[0018] 明显地，文件服务器应用 120 是在操作系统 118 的用户空间中运行以支持网络共享的应用。在该常规配置中，网络接口 120 仅将网络通信从网络 120 传递到操作系统 118，随后操作系统 118 将适当的通信转发到文件服务器 120。进而，文件服务器 120 通过将这些文件系统命令翻译为块级命令从而解释并执行各种网络文件系统命令以便存储在存储驱动器 102 上的网络共享。因此，常规存储架构要求用户空间应用的使用从而执行网络级文件系统命令。

[0019] 如上所述,该实施例克服对支持驱动器上更高级别命令(诸如,网络级文件系统命令)的用户空间应用和软件的需求。在某些实施例中,驱动器可以直接执行文件系统命令,诸如 NFS 命令、SMB 命令等。

[0020] 图 2 显示存储驱动器 200 的示例性实施例,其可以执行文件系统命令。如图所示,存储驱动器 200 包括媒介组件 202、存储媒介 204、控制器 206、存储器 208、文件系统处理器 210、网络接口 212 以及文件系统 214。下面将简单描述这些组件。

[0021] 媒介组件 202 和存储媒介 204 表示由存储驱动器 200 使用以存储和存取数据的组件和物理介质。例如,在一个实施例中,媒介组件 202 和存储媒介 204 可实施为硬盘介质、固态介质、闪存等。任何形式的存储技术和介质都可以被用在实施例中。

[0022] 控制器 206 控制存储驱动器 200 的操作。例如,控制器 206 可以控制对媒介组件 202 以及存储媒介 204 进行数据的读取和写入的操作。控制器 206 可以使用本领域技术人员公知的组件实施。

[0023] 存储器 208 用作控制器 206 的临时存储器位置,诸如用于缓存数据,指令排队等。如图所示,存储器 208 还为至少一部分已经在存储驱动器 200 上实施的文件系统提供存储位置。

[0024] 文件系统处理器 210 表示处理资源,其接收和执行用于存储驱动器的网络级文件系统命令。如下面所述,文件系统处理器 210 可以通过多种方式来实施。文件系统处理器 210 处理对存储驱动器的资源(诸如,媒介组件 202 和存储媒介 204)的特权化存取。特权化存取可以指代对组件的专用组严格的那种形式的存取,或优先或与正常存取请求(诸如,运行在操作系统的用户空间中的常规应用软件请求)不同处理的存取。例如,在某些实施例中,特权化存取可指代不要求命令翻译或中间组件的直接存取。根据某些实施例,内核设备驱动器是具有特权化存取的组件的一个例子。在另一个实施例中,如下面进一步描述的,文件系统处理器 210 可以基于硬件或固件实施。

[0025] 在某些实施例中,存储驱动器 200 是运行嵌入式操作系统的嵌入式设备。在其他实施例中,可以通过操作系统(诸如,LINUX 或 Windows)来加载存储驱动器 200 中的控制器 206,并且文件系统处理器 210 包括一组内核设备驱动器,所述内核设备驱动器经配置以执行网络级文件系统命令,诸如 SMB 命令或 NFS 命令。文件系统处理器 210 还可以包括块级存储媒介设备驱动器(诸如,SATA 驱动器),从而为媒介组件 202 和存储媒介 204 提供接口。

[0026] 在某些实施例中,文件系统处理器 210 可以包括实施各种安全保护的设备驱动器,从而例如防止缓存器溢出和某些操作限制。可以通过已知的编程语言(诸如 C 等)来编写设备驱动器。

[0027] 用于文件系统处理器 210 的内核设备驱动器可以通过各种方式被包括在存储设备 200 中。例如,在一个实施例中,内核设备驱动器可以作为可加载内核模块而被加载到控制器 206 上。可替换地,文件系统处理器 210 可以作为控制器 206 中嵌入式操作系统的一部分来安装。

[0028] 因为文件系统处理器 210 在操作系统的内核空间中操作,因此所述文件系统处理器 210 相对于运行在存储设备 200 上的用户空间中的其他代码或软件可以运行对存储驱动器 200 的资源(诸如,存储器 208、控制器 206、媒介组件 202 以及存储媒介 204)的特权化

存取。

[0029] 文件系统处理器 210 中的这些形式的内核设备驱动器允许存储设备 200 通过增强的特权在内核空间中执行网络文件系统命令, 这比使用用户空间软件更快且更有效。

[0030] 在某些实施例中, 可以使用硬件和 / 或固件组合实施文件系统处理器 210。例如, 在一个实施例中, 文件系统处理器 210 可以包括固件 ( 诸如, 现场可编程门阵列, 或专用集成电路 ), 其可经配置以解释和执行网络级文件系统命令。在一个实施例中, 文件系统处理器 210 包括经配置以执行 SMB 命令的固件。在另一个实施例中, 文件系统处理器 210 包括经配置以执行 NFS 命令的固件。如上所述, 文件系统处理器 210 可以经配置以执行任何类型的文件系统或协议。

[0031] 网络接口 122 提供存储驱动器 100 的通信接口。网络接口 122 可以支持任何类型的有线或无线通信, 诸如以太网、千兆以太网、无线带宽、Wi-Fi 等。

[0032] 在一个实施例中, 网络接口 122 包括一个或多个用于发送和接收包的硬件或固件组件。例如, 网络接口 122 可以经配置以封装 / 解封装传输 (transport) 各种网络级文件系统命令 ( 诸如, SMB 或 NFS 命令 ) 的 TCP/IP 包。随后网络接口 122 将这些命令传递到文件系统处理器 210 以便解释和执行。

[0033] 在一个实施例中, 网络接口 122 还可以提供各种安全特征, 诸如加密 / 解密。例如, 网络接口 122 可以包括加密芯片或系统。

[0034] 在某些实施例中, 网络接口 122 由在控制器 206 和 / 或文件系统处理器 210 上执行的设备驱动器支持。在一个实施例中, 网络接口 122 的设备驱动器在由控制器 206 上运行的操作系统提供的内核空间中运行。

[0035] 图 3 示出一个实施例的示例性流程。为了以说明为目的, 图 3 涉及经由网络接口 ( 诸如, 以太网或千兆以太网接口 ) 附接到网络的存储驱动器 200, 并且涉及基于从客户端 ( 未示出 ) 或其它主机设备 ( 未示出 ) 接收的 NFS 命令提供对其文件的存取。当然, 其它类型的网络级文件系统命令可以由存储驱动器 200 处理和执行。

[0036] 如图所示, 在阶段 300, 存储驱动器 200 可经由网络接口 212 接收一个或多个网络文件系统 (“NFS”) 命令。NFS 命令可以被封装在一个或多个数据包 ( 诸如 IP 数据包 ) 中, 并且可以经由网络上的以太网帧来传输, 其中网络接口 212 被连接到所述网络。如上所述, 网络接口 212 可以是以太网或千兆以太网接口。用于这种类型的接口的组件是本领域技术人员已知的。网络接口 212 随后可以执行涉及解封装来自以太网帧和 IP 数据包的数据包的各种数据接收和传输功能。

[0037] 在阶段 302, 网络接口 212 可以将 NFS 命令转发到文件系统处理器 210。文件系统处理器 210 可以包含硬件和 / 或在控制器 206 上运行的软件, 所述文件系统处理器 210 具有对存储驱动器 200 的其它组件的特权化存取。如上所述, 例如, 在一个实施例中, 文件系统处理器 210 可以被实施为存储驱动器 200 上特殊目的固件 ( 诸如, 现场可编程门阵列或 “FPGA”、专用集成电路、或 “ASIC”)。可替换地, 文件系统处理器 210 可以是程序代码, 诸如在控制器 206 上运行的软件和设备驱动器的集合。

[0038] 为了支持网络级文件系统命令 ( 诸如 NFS 命令 ) 的执行, 文件系统处理器 210 可以存取存储在存储器 208 中的文件系统数据 214。文件系统数据 214 还可以通过文件系统处理器 210 从存储媒介 204 取回。

[0039] 在阶段 304, 文件系统处理器 210 通过使用其对存储驱动器 200 的资源的特权化存取从而执行网络级文件系统命令, 诸如 NFS 命令或 SMB 命令。值得注意, 与常规存储驱动器相比, 存储驱动器 200 自身能够执行文件系统命令, 而非依靠另一设备, 诸如运行应用程序的客户端计算机或在驱动器上运行的应用级软件。相反, 如图 2 和 3 所示, 存储驱动器 200 可以直接接收文件系统命令并使用文件系统处理器 210 执行这些命令。

[0040] 文件系统处理器 210 可以基于特权化存取执行文件系统命令。例如, 文件系统处理器 210 可以是连接到控制器 206 并被给予对其处理逻辑存取的固件, 并可以对媒介组件 202 和存储媒介 204 具有特权化存取。此外, 文件系统处理器 210 可以对存储器 208 具有特权化存取。

[0041] 可替换地, 在另一个实施例中, 文件系统处理器 210 可以是在控制器 206 上运行的软件, 不过作为内核级应用或在安装在控制器 206 上的操作系统内运行的内核驱动器而具有特权化存取。

[0042] 在一个实施例中, 文件系统处理器 210 通过提供文件级命令, 而非块级命令对控制器 206 和媒介组件 202 执行文件系统命令。为了解释文件级命令, 控制器 206 和媒介组件 202 可经配置以采用将文件映射到存储媒介 204 上位置的映射或数据结构。例如, 下面示出用于执行 NFS 命令的示例表。

[0043]

文件操控	节点	生成	存储媒介领域

[0044] 这样的表格可由文件系统处理器 210 维持并存储在存储器 208 中或存储媒介 204 上的位置中。上面的表格仅是示例性的, 且本领域技术人员将理解这些实施例可以为其它类型的文件系统和文件系统命令实施。

[0045] 上面公开的特定实施例的特征和属性可以通过不同方式组合以形成额外实施例, 所有这些实施例都落入本申请的范围。虽然本公开提供某些实施例和应用, 然而对本领域技术人员明显的其它实施例, 包括不提供这里所描述的所有特征和优点的实施例, 也都在本公开的范围。因此, 本公开的保护范围仅由随附权利要求的参考限定。



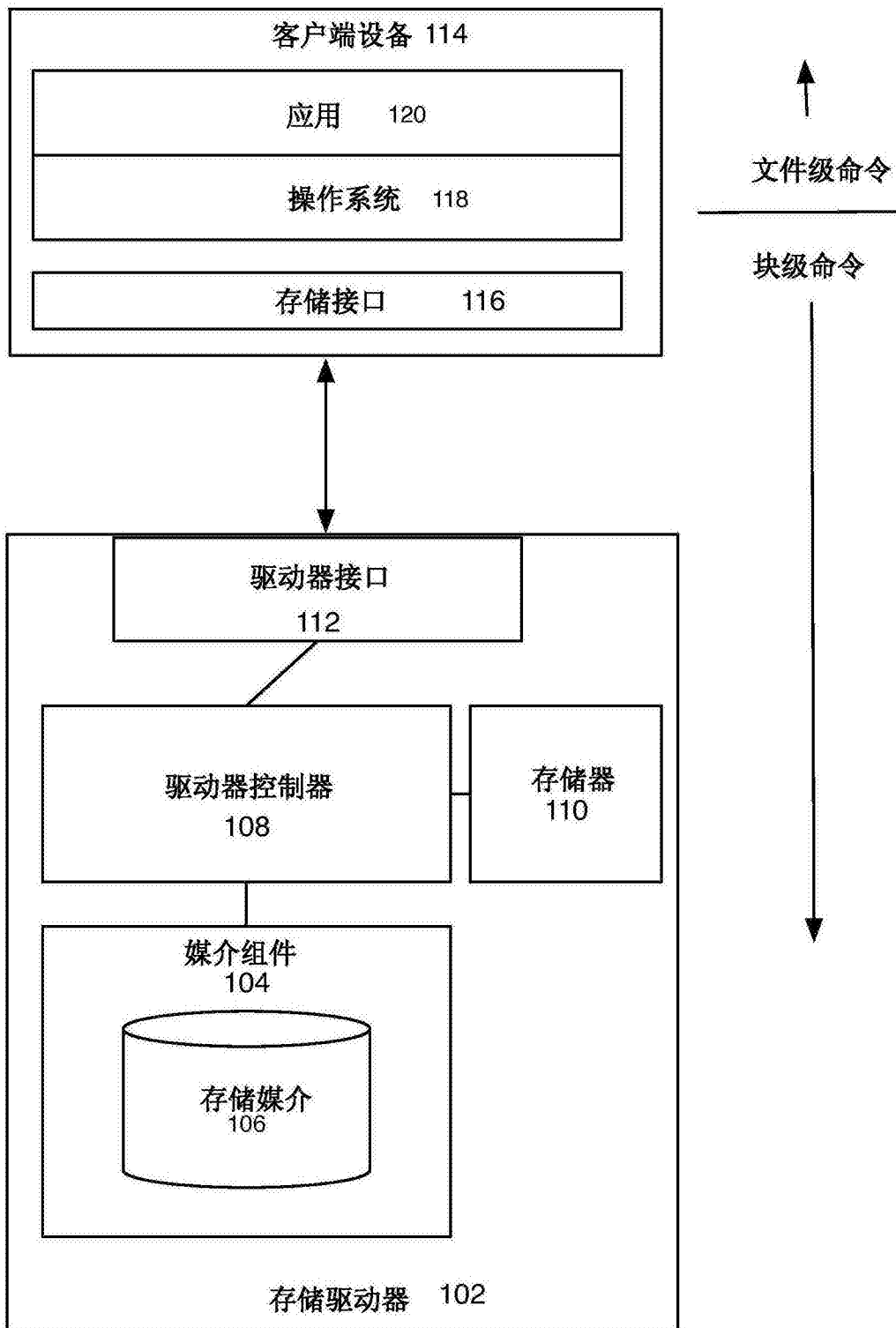


图 1A(现有技术)

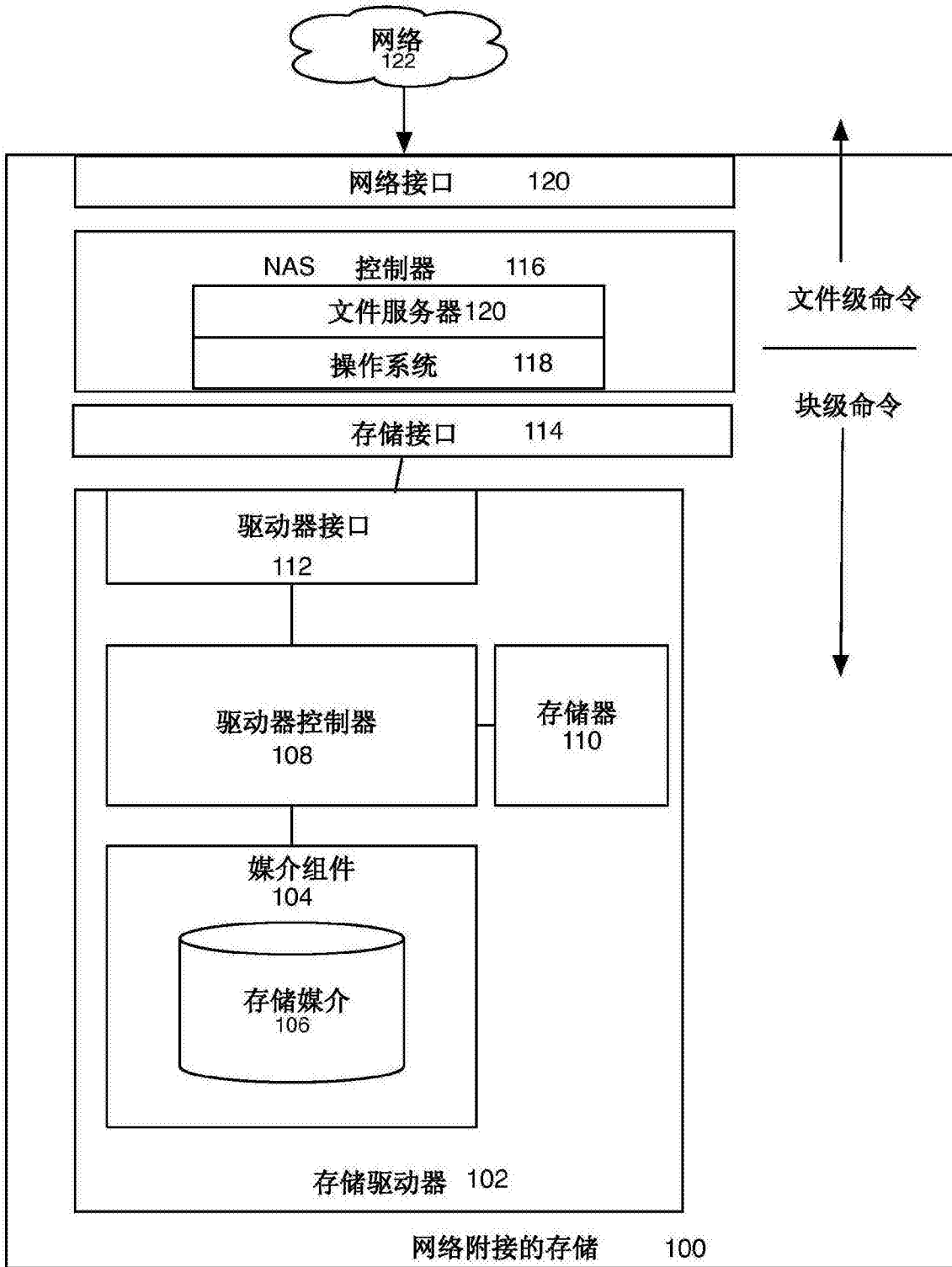


图 1B(现有技术)

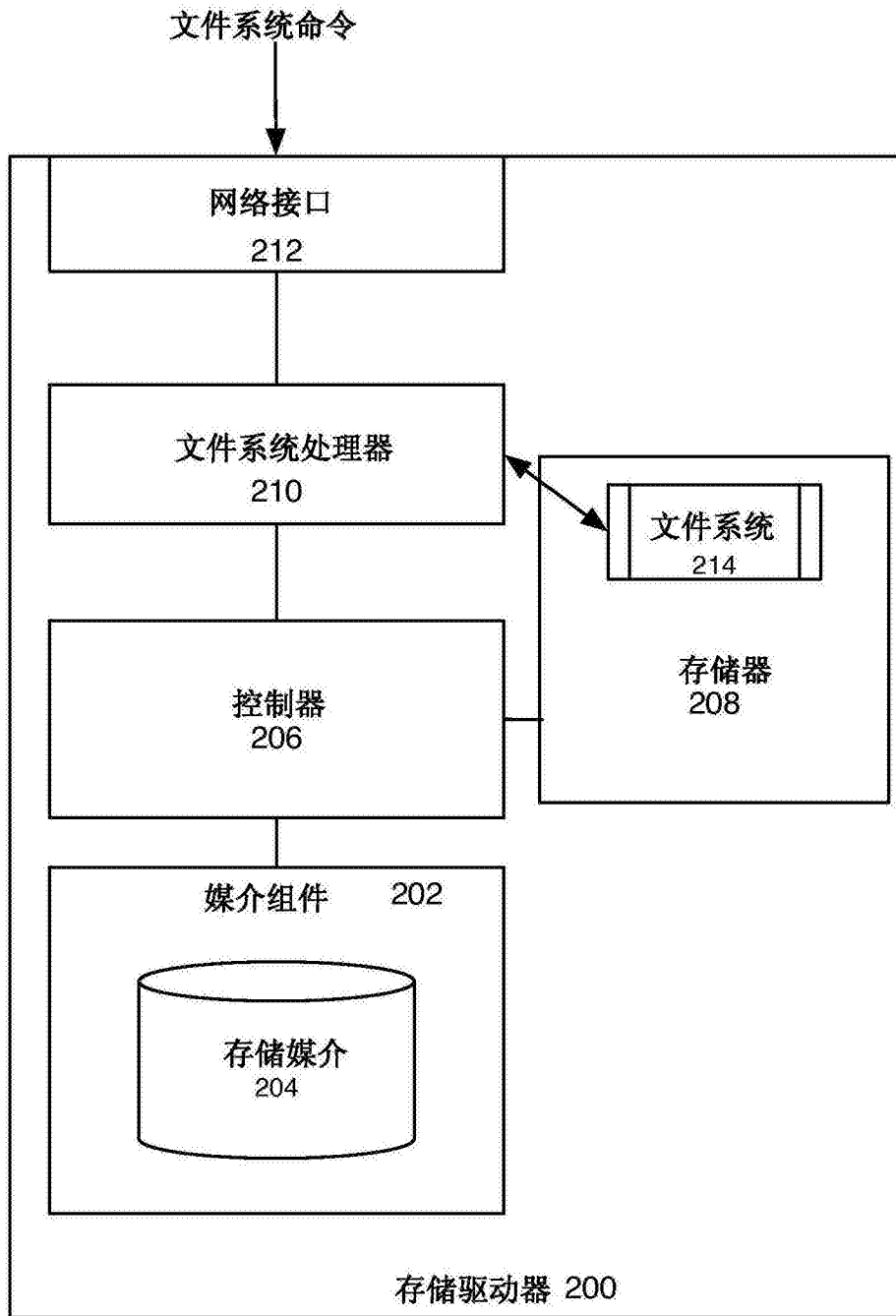


图 2



图 3