



(12) 发明专利

(10) 授权公告号 CN 101243413 B

(45) 授权公告日 2013.08.14

(21) 申请号 200680030218.5

(51) Int. Cl.

(22) 申请日 2006.06.23

G06F 12/00(2006.01)

(30) 优先权数据

(56) 对比文件

60/693,715 2005.06.24 US

WO 9309494 A1, 1993.05.13, 全文.

(85) PCT申请进入国家阶段日

US 20030191911 A1, 2003.10.09,

2008.02.19

US 6795966 B1, 2004.09.21,

(86) PCT申请的申请数据

US 20040172574 A1, 2004.09.02, 全文.

PCT/US2006/024479 2006.06.23

审查员 于春晖

(87) PCT申请的公布数据

W02007/002398 EN 2007.01.04

(73) 专利权人 信科索尔特公司

地址 美国新泽西

(72) 发明人 彼得·志雄·刘 苏比·阿卡尔亚

(74) 专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 赵科

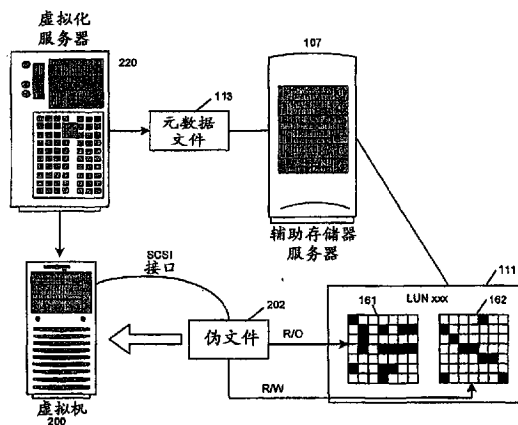
权利要求书2页 说明书6页 附图1页

(54) 发明名称

用于对备份映像进行虚拟化的系统和方法

(57) 摘要

当进行备份时,使用由备份软件创建的映像来重新创建整个机器(好像其处于过去的某个时间点)的一种工具。该工具可以被延伸,以便引出一组机器,其一起起到服务于一些逻辑商业功能的作用,正如处于集群中或联合服务器中,并且进一步延伸,这样整个数据中心可以从备份映像被虚拟化。该提供的虚拟化服务器可以在灾难情况下起到备用数据中心的作用,或者满足维护窗口的作用,以便起到低成本瞬时灾难恢复的作用。一组虚拟机可以代替物理机器一段时间,然后经由裸机恢复和重新同步的组合从形成虚拟机盘的正在运行的 LXJN 中重新同步或再播种物理机器。



1. 一种用于根据物理机器的备份映像虚拟化所述物理机器的方法,包括:
 - a) 以用于虚拟化服务器的虚拟机映像的格式创建所述物理机器的备份映像,还包括与
所述备份映像一起、对所述备份映像补充地保存机器配置信息 ;和
 - b) 在所述虚拟化服务器下加载所述备份映像作为虚拟机映像。
2. 根据权利要求 1 的方法,其中所述备份映像是所述物理机器的时间点快照。
3. 根据权利要求 1 的方法,其中所述创建所述物理机器的所述备份映像的步骤包括:
 - i. 获取所述物理机器的基础级别时间点快照 ;
 - ii. 在存储设备上存储时间点快照映像 ;
 - iii. 获取所述物理机器的块级别的增量时间点快照,其中增量时间点快照仅表示自从
先前时间点快照以来已经变化的块 ;
 - iv. 以能够被直接用作安装的存储器单元的方式构造增量时间点快照 ;
 - v. 与所述时间点快照映像一起保存表示所述物理机器在每个快照映像的时间点时的
状态的信息。
4. 根据权利要求 1 的方法,其中所述加载所述备份映像的步骤包括:
 - i. 从最近的时间点快照或任何其他时间点中选择表示所述物理机器的一个节点或一
组节点 ;
 - ii. 通过创建表现所述备份映像的顶部的包络的伪文件,将所述备份映像变换成 LUN ;
 - iii. 从与所述备份映像结合表现的文件中读取元数据,并使用所述元数据来发布创建
虚拟机的指示 ;和
 - iv. 从所述 LUN 引导所述虚拟机。
5. 根据权利要求 4 的方法,还包括通过分别写入新创建的数据和 / 或变化后的数据来
维持所述备份映像的完整性。
6. 根据权利要求 4 的方法,还包括以适于操作系统的引导记录填充所述 LUN,并创建分
区表跳转,并在需要时修复文件系统的引导扇区。
7. 根据权利要求 4 的方法,还包括修改在所述 LUN 上的数据和 / 或配置,使得从所述
LUN 引导能够成功。
8. 根据权利要求 1 的方法,还包括:
 - c) 在所述虚拟化服务器下作为虚拟机运行所述虚拟机映像以处理新的输入 ;
 - d) 在创建所述备份映像之后的一个时间重启所述物理机器 ;和
 - e) 将所述物理机器上的数据与所述虚拟机上的数据同步。
9. 根据权利要求 4 的方法,还包括将后台中的所有块从所述备份映像复制到所述 LUN。
10. 一种用于虚拟化多个物理机器的方法,包括:
 - a) 以用于虚拟化服务器的虚拟机映像的格式创建每个所述物理机器的备份映像,还包
括与所述备份映像一起、对所述备份映像补充地保存相应的物理机器的机器配置信息 ;和
 - b) 在所述虚拟化服务器下加载所述备份映像作为多个虚拟机映像。
11. 根据权利要求 10 的方法,其中,
所述步骤 a) 还包括:
创建每个所述备份映像作为相应物理机器的时间点快照 ;和
与每个这样的快照映像一起保存表示相应物理机器在这样的快照映像的时间点处的

状态的信息,其中这样的状态包括所述物理机器的网络连接,以及,

所述步骤 b) 还包括:

与所述备份映像一起加载所述保存的有关网络连接的信息,并使用这样的信息来在所述虚拟化服务器内重建每个这样的物理机器的网络连接。

12. 根据权利要求 11 的方法,其中所述多个物理机器表示数据中心中的计算机,并且所述方法被用于创建数据中心的虚拟化。

用于对备份映像进行虚拟化的系统和方法

[0001] 相关申请的交叉参考

[0002] 本申请要求于 2005 年 6 月 24 日提交的美国临时申请 NO. 60/693,715 的申请日的优先权,该申请在这里引为参考。本申请还通过参考而结合了与本申请同一天提交的、申请号为 NO. ____、题为“System And Method for High Performance Enterprise DataProtection”的共同发明且共同受让的申请的整个公开内容。

技术领域

[0003] 本发明属于信息技术领域,更具体而言,本发明涉及高性能、企业级的备份和灾难恢复系统,以及虚拟化技术在该领域中的使用以便提供更多的优点和益处。

背景技术

[0004] 备份和灾难恢复技术领域中的现有技术在上述参考和结合的共同专利申请中列出的参考文献中有所描述。

[0005] 提供虚拟化服务的系统的实例例如包括:VMWare(www.vmware.com);Microsoft Virtual Server(www.microsoft.com/windowserversystem/virtualserver/default.msp);和 Xen(www.xensource.com)。

[0006] 在虚拟化技术的现有技术中,在企业级的备份和灾难恢复操作中已经有了一些应用。

[0007] 例如,广泛使用的虚拟化服务器、诸如上述那些虚拟化服务器都能够通过创建新的(空的)虚拟机“外壳(shell)”;在新的虚拟机中安装操作系统;将应用程序安装到新的虚拟系统中;并且为那些应用程序复制用于其他虚拟或真实的(即,物理的)机器的现有数据,从而创建和保存虚拟机。

[0008] 此外,已知在虚拟机上安装备份客户端软件并向虚拟机恢复从真实机器创建的备份集,包括处理器和外围设备的状态。参见 Traut,美国专利公开 No. US 2006/085792 A1(2006 年 4 月 20 日)。但是,这与恢复到真实机器相比,操作上没有区别。

[0009] 此外,虚拟化软件可以提供从第三方备份格式导入数据的能力。参见“Importing Virtual Machines and System Images from OtherFormats”(www.vmware.com/support/ws55/doc/ws_newguest_vm_importer.html)。同样,这与当在向真实机器恢复数据中导入外部备份格式时可能执行的“导入”操作基本上是相同的类型。

[0010] 现有技术并没有提供任何以如下方式来备份真实机器的过程,即整个机器可以在虚拟服务器控制之下在虚拟化环境中从备份映像中被重新创建为其在过去特定时间点所存在的那样。

发明内容

[0011] 本发明的一个目的就是提供一种设备,用于利用备份软件所创建的映像(或备份软件所重新构成的其他软件所创建的映像)来将整个机器重新创建为其在过去进行备份

时那样。

[0012] 本发明的另一目的是扩展这样的设备,以构成如集群或联合服务器中那样的一起提供一些逻辑业务功能的一组机器。

[0013] 本发明的再一目的是扩展上述方法,使得可以由备份映像虚拟化整个数据中心。

[0014] 本发明的还一目的是提供一种设备,用于将根据备份所创建的虚拟机重新与相应的物理机器同步。

[0015] 本发明所提供的虚拟化服务器可以用作在灾难情况下保持不变的替换数据中心,或用于满足维护窗口从而实现低成本即时灾难恢复。一组虚拟机可以代替物理机器一段时间,然后经由裸机恢复 (baremetal recovery) 和从形成虚拟机盘的活动 LUN 重新同步的组合而被重新同步或重新播种 (re-seed) 到物理机器。

[0016] 在一个实施例中,根据期望,为一个服务器或任意更多数量的服务器,通过创建虚拟化软件所期望的配置和映像格式的备份映像,并在虚拟化软件中重启这些机器,从而能够提供如上所述的设备。因为没有涉及数据移动或传统的恢复,因此这是接近即时的过程。虚拟机实际上由读写 LUN 支持,其中 LUN 由备份软件创建,备份软件利用仅用于变化数据的附加空间来清除只读备份映像。基本的操作系统软件也被适当地修改,从而可以进行从物理硬件到虚拟硬件的过渡。

[0017] 本发明的一个实施例正被实施为 Backup Express[®] (BEX) 的一部分,该 BEX 是 Syncsort 公司的软件产品,Syncsort 公司是本申请的受让人。

[0018] 本发明的其他目的和优点将通过附图和以下详细描述而变得显而易见。

附图说明

[0019] 图 1 是示意性地表示本发明所使用的创建备份映像的框图。

[0020] 图 2 是示意性地表示使用由图 1 的操作所产生的备份映像来启动虚拟化服务器之下的虚拟机的框图。

具体实施方式

[0021] 以下将描述本发明各方面的几个优选实施例,以显示系统如何被构成以执行本发明的细节,以及可以被用于使用这样的系统并实现这样的方法的步骤。这些实施例仅仅是示例性的,并且本发明并不限于所示的特定实施例。例如,某些优选实施例是针对具有特定存储器硬件、操作系统和虚拟化软件的实现来描述的,但是应当理解,以下的公开是为了使本领域技术人员能够易于将所陈述的内容教导应用到其他存储器硬件、操作系统和虚拟化环境中。任何具体实施例的特定特征不应当被理解为限制所要求保护的范围。

[0022] 定义

[0023] 在本申请中,以下术语具有定义的含义:

[0024] 备份软件:创建捕捉增量变化并在辅助存储器上保存过去时间点的备份映像。备份软件创建应用一致映像,并附加地捕捉机器配置,包括持久状态和不稳定状态。

[0025] 辅助存储器:与主存储器(产品数据驻留在主存储器中)相区别,这是备份的目的地,也是形成虚拟机盘的 LUN 的数据储存库。只有变化需要附加存储器,因此除了备份所必需的之外,只需要很少的辅助存储器。该存储器可以是一次写多次读 (WORM) 的,以便支持

不能改变的内容保持,以满足法律要求。

[0026] 虚拟化硬件:在要求时运行虚拟化软件和虚拟机的单个机器或一组机器。依赖于配置的相同硬件也可以用作备份目的地,并运行备份智能。

[0027] 虚拟化软件:对物理硬件虚拟化,从而允许多个虚拟机在给定的硬件上运行。每个虚拟机在虚拟硬件上运行。操作系统与物理硬件去耦。

[0028] 卷:备份的单位,包括在块级别上备份的多个文件和目录的单个文件系统。

[0029] 备份机制

[0030] 通过传输为第一备份分配的块、然后后续备份的变化的块,创建备份映像。为一组卷创建与应用一致的快照,所得到的数据被一致地移动到辅助存储器。在第一备份之后,只有自从最后备份以来的变化被捕捉并被应用到现有辅助映像中。与数据一起,机器配置和状态被传输到辅助存储器,并且被保存作为用于给定机器的元数据。对于备份机器的每个源卷存在单独备份映像。这些映像以及元数据的快照保存该时间点状态,以用于将来的虚拟化。被捕捉的元数据包括硬盘驱动配置(包括几何形状、尺寸、逻辑卷特征等等)、以及存储器、和网络控制器细节。快照捕捉足够的数据和状态,以在以后某个时间根据需求虚拟化整个机器。以这种方式创建的备份映像被保存为特定虚拟化服务器所要求的格式。

[0031] 图 1 显示了该过程。主系统 100 被备份到备份软件所创建的读写 LUN 111,其中备份软件利用仅用于变化数据的附加空间 162 清除只读备份映像 161。作为过程一部分被捕捉的元数据被保存在元数据文件 113 中。

[0032] 在某些情况下,备份映像可以预先存在或者作为 LUN 被外部地创建在 SAN 或 NAS 存储器上复制卷上。在该情况下,备份软件需要靠着这些 LUN(或映像)运行,以产生适合于虚拟化的元数据和封装机器状态。

[0033] 备份存储器也可以是 WORM(一次写多次读)的,以满足规定的或严格的要求。

[0034] 虚拟化步骤

[0035] 当用户希望时,这些备份映像可以在非定制(of-the-shelf)虚拟化软件的帮助下能够“显得逼真”或被虚拟化。因为用于机器的备份映像以虚拟化软件所要求的形式被保存,因此可以直接从备份映像重启虚拟机,以将机器状态恢复到备份时间点(通常是最接近的备份时间点)。因为没有数据移动,因此整个过程非常快。总时间完全由虚拟机引导所花费的时间决定。

[0036] 图 2 和以下的描述显示了该过程如何工作,在一个实施例中:

[0037] 用户从最近的快照或任何其他时间点中选择表示各个机器、集群、或构成重要数据中心功能的一组机器的一个节点(或一组节点)。(为了简便起见,图 2 假设所进行的选择是用于单个节点的,在该情况下,图 1 的主系统 100。)如果虚拟化的需求是用于验证或是为了实现其他轻量分析功能,则这些机器可以出现在它们自己的私用网络中,或者它们可以出现在公共企业网络上以取代故障机器。软件根据目标硬件上可用的物理硬件总量自动选择 CPU 数量和存储器的缺省值,或者用户可以定制这些值。选择、虚拟化、然后满足的过程可以是反复的、改进的、记住的和列举的。来自多个运行的模板可以被保存,并且当需要时立即被应用。

[0038] 一旦完成上述选择,创建虚拟机 200 的适合于虚拟化软件的虚拟机配置,然后将其加载在虚拟化服务器 220 上。基于元数据文件 113 的内容,创建对 CPU 数量、存储器量、

网络类型和 SCSI 控制器以及硬盘驱动器数量的指示,作为配置的一部分。

[0039] 通过创建表示备份映像 111 顶上的包络的伪文件 202,源卷的每个备份映像被转换成 LUN。该 LUN 被虚拟机看作是正常读 / 写 SCSI 硬盘驱动。向 LUN 写并不影响备份映像 161 的完整性,但是被单独地白痴 162(尤其是如果辅助存储器是 WORM 时,则由于技术和商业原因,写是不可能的)。这些 LUN 经历再引导,并且保留对其的所有写入。

[0040] 根据源机器是单机还是集群的,LUN 被不同地创建。对于单机机器,由每一卷创建单分区 LUN,而不管在源上存在多少物理硬盘驱动。逻辑卷管理器控制的卷被剥夺该特性。(通过虚拟化盒上基本存储器的特性实现对于虚拟盘的性能和 / 或可靠性标准)。在集群的情况下,共享硬盘驱动的准确数量被重建为一组多分区 LUN,使得集群软件可以工作。

[0041] 需要注意的重要事项在于,实现此不需要附加的存储器。只有新创建的和 / 或改变后的数据需要新的存储器 (162)。(这通常是原始存储器的 5%)。不仅仅不需要附加硬件来虚拟化多个机器,而且几乎不需要附加存储器来实现此。(通常,辅助存储器上可用的额外容量通常就足够了。)

[0042] 注意:可以根据同一组分别具有其自己的独立生存周期的备份映像创建多个虚拟机,从而原始映像没有改变。

[0043] 4. 为了虚拟机可以引导,引导和系统驱动被灌注以适合于操作系统的引导记录,以及分区表跳转被创建,并且在需要时文件系统的引导扇区被修复 (fix)。这使得虚拟机 BIOS 能够将控制转移到 LUN 上的引导码 (boot strap code),使得备份机器上的操作系统可以被启动,这又引出数据库、商业营运应用等等。(该步骤特定于正被虚拟化的操作系统映像)。

[0044] 5. 然后,使目标 VM 的系统驱动对于虚拟化盒可视,并且被称为 Osfixer 的过程运行,其修改系统驱动上的文件和 / 或配置,使得前一引导到虚拟硬件的步骤可以成功。这涉及依赖于源和目标的 OS 特定步骤,以重新实现用于虚拟机引导的系统驱动。

[0045] 6. 然后,一个或多个虚拟机被启动,以完成该过程。根据这些机器为何被引出,这些机器可以出现在公共网络上或者出现在隔离的网络上。接下来发生什么就取决于需要被实现的业务功能,并且将在以下介绍。

[0046] 商业可应用性

[0047] A. 替代站点故障或病毒攻击或其他反复破坏 (rollingcorruption):在移动数据中心或甚至在一些技术理解基础中,在共享设备处,由于业务连贯性的原因,几乎不利用除了已分配给盘备份之外的附加存储器地,临时性地在备用站点上以缩减的硬件根据备份映像虚拟化重要的机器组(包括集群)。在病毒攻击或反复破坏的情况下,时间点选择不必是最近的,但是必须在攻击或破坏之前。在存在反复或蔓延的破坏的情况下,各个机器可以退回到各个时间点。运转中断的停机时间可以低至几分钟。

[0048] B. 规定符合性:可以以最小的硬件成本从过去的对于规定重要的时间点来虚拟化机器或一组机器这一事实促进了符合性。不仅仅所有数据可以是可用的,而且具有相关应用的整个机器可以全部被恢复,从而在面对审计时表现商业部分的根本诚意。法律发现、符合性,和 / 或审计过程可以从几个月缩短为几天,甚至几个小时。

[0049] C. DR 训练 (DR Drill) / 备份验证:由于花费和时间限制,如今大多备份未经验证。DR 训练也是耗时且难忍受的。备份虚拟化使其变得简单、高效、快速,并且商业必须具有。

这可以被用作例行程序业务实践,作为业务数据和配置的更有效的保护装置。

[0050] D. 与后代遗留数据一起的近线化遗留应用 (Near-lining legacy application): 通常,备份仅保护数据,但是备份虚拟化保护可以在虚拟机上被重新创建并在创建该应用程序的公司破产或停止支持很久之后能够被遗留应用程序理解的数据和配置。(如果备份已经失效,则备份虚拟化还可以从磁带中实现,尽管这需从磁带恢复到磁盘。)

[0051] 数据仓库和通用目的假设分析:当基于数据库的商业营运应用需要输出用于季度预告、分析、报告、测试等的的数据时,招致非常大量的附加存储器和和管理开销(更不用说在高薪聘请的分析师等待期间所损失的创造性)。备份虚拟化不仅使数据可用,而且机器和主机数据库可以以小部分硬件和存储器成本并且近似即时地从期望的时间点(分析上合适的)重新创建。对商业的成本节省是巨大的。

[0052] 从物理机器到虚拟机的迁移:一旦从备份映像虚拟化物理机器,这些就可变得持久,从而允许从遗留硬件进行迁移。

[0053] 从一个机器克隆多个机器:因为可以从给定时间点的单组映像中创建多个虚拟机,因此这允许很容易地克隆“好”配置以及多个用户使用多个虚拟机,其中每个虚拟机共享同一组映像并使用仅用于在各个虚拟机上进行的实际变化的存储器。

[0054] 后虚拟化

[0055] 根据为什么使用备份虚拟化,主要有三种情况。

[0056] 1、需求是短期的:在该情况下,在“备份到虚拟(B2V)”已经起到商业目的之后,其可以被拆卸和忘记,从而释放资源和存储器空间。该组备份映像(直到它们失效)以它们原来形式保持可用,以在需要时用于进一步的虚拟化。

[0057] 2、需求是长期的:虚拟机可以被证明是有用的,并且需要更多力量。这可以通过从它们的备份快照中释放基本 LUN,然后移动它们和/或赋予它们更多虚拟硬件来实现。这受到将后台中的所有块从备份映像复制到 LUN 的影响。另一选择是经由裸机恢复、然后将其与活动的虚拟机同步,以最小的停机时间从同一时间点备份映像中播种(seed)物理机器。

[0058] 3、需求是短期的,但是需要生存周期结束同步:当备份虚拟化被用于代替故障机器或在预定维护期间,在未来某个时间点的重新同步是必需的。这是因为,虚拟机在该时间段期间已经是活动的,并且数据已经从备份映像的时间点脱离。在原始故障机器已经经由裸机恢复而被恢复之后或者在预定维护完成之后,如今存在于虚拟机中的新的状态和数据需要被重新同步。通过保持虚拟机运行直到它们准备好被切换,重新同步过程使应用程序停机时间最小化。以下将介绍如何实现。

[0059] 虚拟机如何与真实机器会聚:作为虚拟机虚拟硬盘基础的 LUN 跟踪经由虚拟机的写操作,并将其独立于备份映像进行存储。在 LUN 上变化的数据在稍后时间(由用户指示)或许需要重新同步的情况下,要求这些 LUN 以频繁时间间隔被检查校验点。这包括当发生校验点时将一个时间段中的所有变化保存到被标记的文件中,然后切换到新的校验点文件。每个卷的每个 LUN 具有一系列包含校验点之间变化的校验点文件。重新同步过程开始于裸机过程的结束,或者由用户启动,并且将数据从这些校验点文件复制到用于该卷的现有或恢复的数据的顶上。

[0060] 如果 OSFixer 已经改变了系统驱动以使其能够在虚拟硬件环境下可引导,则该

特定校验点文件被忽略用于重新同步。在目标硬件不同于源硬件的情况下,需要附加的 OSFixer 步骤以使映像符合目标硬件。一旦所有校验点文件被复制,虚拟机就被停止,并且最后的校验点文件被应用。真实机器然后接管。于是,虚拟机可以被是循环的、再生的资源。

[0061] 很明显,这里所描述的实施例实现了本发明所陈述的目标。虽然当前优选实施例已经被详细地描述,但是对于本领域技术人员来说,本发明的原理可以由其他设备、系统和方法来实现,而不背离如下面权利要求所定义的本发明的范围和实质。

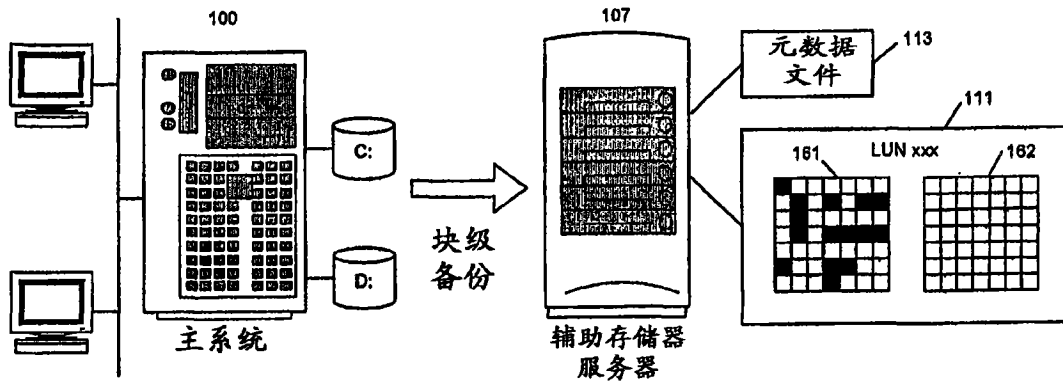


图 1

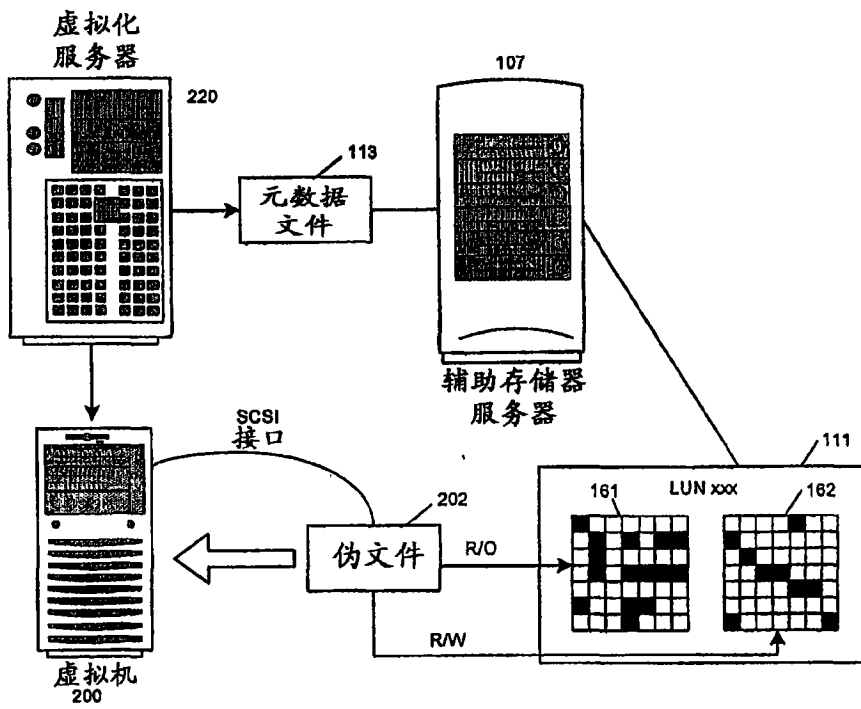


图 2