



(12)发明专利

(10)授权公告号 CN 104572248 B

(45)授权公告日 2018.01.23

(21)申请号 201410790543.8

G06F 11/14(2006.01)

(22)申请日 2014.12.17

(56)对比文件

(65)同一申请的已公布的文献号  
申请公布号 CN 104572248 A

CN 101609419 A, 2009.12.23,  
CN 103678045 A, 2014.03.26,  
US 2011258385 A1, 2011.10.20,  
CN 101464812 A, 2009.06.24,

(43)申请公布日 2015.04.29

(73)专利权人 中电科华云信息技术有限公司  
地址 200231 上海市徐汇区华泾路509号7  
幢552室

审查员 李晓晖

(72)发明人 崔昊之 丁星 武静

(74)专利代理机构 上海汉声知识产权代理有限  
公司 31236

代理人 郭国中

(51)Int.Cl.

G06F 9/455(2006.01)

G06F 3/06(2006.01)

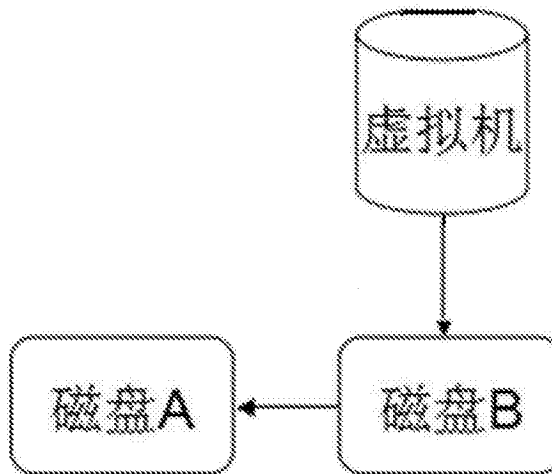
权利要求书1页 说明书3页 附图1页

(54)发明名称

虚拟机动态保存快照的方法

(57)摘要

本发明提供了一种虚拟机动态保存快照的方法,其包括以下步骤:步骤一,使用一特定结构初始化文件,该文件在后续步骤用于分段保存虚拟机状态数据与内存数据;步骤二,创建一块新的虚拟磁盘文件;步骤三,遍历该虚拟机中挂载的所有虚拟设备;步骤四,循环迭代的保存虚拟设备状态与虚拟机内存中变化的部分至指定的文件,并统计每次保存数据量与保存消耗时间;步骤五,每次循环迭代开始,首先统计所有待决的脏数据量,跳出迭代后即加锁虚拟机IO线程;步骤六,解锁虚拟机IO线程,并恢复虚拟CPU线程。本发明在不影响虚拟机内应用服务的前提下完成虚拟机快照操作,并能在需要时根据快照保存的档案完全恢复虚拟机磁盘与内存的当时状态。



1. 一种虚拟机动态保存快照的方法,其特征在于,所述虚拟机动态保存快照的方法包括以下步骤:

步骤一,使用一特定结构初始化文件,该文件在后续步骤用于分段保存虚拟机状态数据与内存数据,此步骤需要初始化该文件的标准I/O流对象,并注册基础读写操作集;

步骤二,创建一块新的虚拟磁盘文件,在虚拟机主进程中注册其结构化数据,并将其置于虚拟磁盘链表的末端;

步骤三,遍历该虚拟机中挂载的所有虚拟设备,使用不同虚拟设备注册的特定方法保存全量的虚拟设备状态数据与客户机内存数据至步骤一创建的文件中;在数据写入的开始、循环迭代及结束过程中均使用魔数分割数据片段,用于读取时识别分段恢复;

步骤四,循环迭代的保存虚拟设备状态与虚拟机内存中变化的部分至指定的文件,并统计每次保存数据量与保存消耗时间;

步骤五,每次循环迭代开始,首先统计所有待决的脏数据量,跳出迭代后即加锁虚拟机IO线程,停止虚拟CPU线程,使动态的数据不再变化,安全完成最后一部分保存动作,完成虚拟机内存与虚拟设备数据的保存后,马上切换虚拟磁盘指针至新的虚拟磁盘,使后续的I/O流向新磁盘设备,而之前的磁盘设备移动为磁盘链表的倒数第二项,同时作为磁盘快照节点,在需要恢复快照时,直接使用该节点生成虚拟机即可;

步骤六,解锁虚拟机IO线程,并恢复虚拟CPU线程,至此虚拟机恢复运行态。

2. 根据权利要求1所述的虚拟机动态保存快照的方法,其特征在于,所述虚拟机动态保存快照的方法涉及的数据对象分为三个部分:虚拟设备状态、虚拟机内存和虚拟机磁盘。

3. 根据权利要求1所述的虚拟机动态保存快照的方法,其特征在于,所述步骤三的魔数分割数据片段包括以下步骤:将数据保存到文件的步骤分为预备、迭代和结尾三部分,文件前两字节依次保存快照动作整体魔数和快照版本信息,接下来排列预备段信息,依次写入预备段开始魔数、单个设备预备快照控制信息、单个设备预备快照数据信息;完成预备段数据后,循环的写入迭代段开始魔数、单个设备迭代快照控制信息、单个设备迭代快照数据信息;最后写入完成段开始魔数、单个设备完成快照控制信息、单个设备完成快照数据信息。

## 虚拟机动态保存快照的方法

### 技术领域

[0001] 本发明涉及一种保存快照的方法,具体地,涉及一种虚拟机动态保存快照的方法。

### 背景技术

[0002] 如今,云计算技术越来越广泛的得到使用,正在很多领域逐渐取代传统的IT服务模式,为IT行业的市场带来巨大的变革,也为IT技术的创新开辟了新的方向;虚拟化是构建云基础架构不可或缺的关键技术之一,对比传统的物理资源,虚拟化技术在资源的有效利用、动态调配和高可靠性方面有着巨大的优势;QEMU-KVM虚拟化(简称KVM)作为一种新兴的开源虚拟化软件,因其虚拟化效率高、与Linux内核完美集成、操作配置简单等优势,与老牌的虚拟化软件如VMWARE、Xen等一起,并列为现今主流的虚拟化技术,并有逐步取代的趋势;但是KVM虚拟化因其发展时间较短,在基本功能上仍有一些欠缺之处,比如:对于正在运行的虚拟机,无法动态的、集成的备份其内存与磁盘的内容(虚拟化快照技术),一种传统方法快照过程中会主动挂起虚拟机服务,快照保存流程完成后再恢复虚拟机服务。另一种传统方法只能动态保存虚拟机磁盘范围的内容,所以在做快照恢复操作时无法还原到虚拟机快照时间点的运行中状态;所以,在使用快照或快照恢复功能时,都会在很大程度上影响云平台的服务质量。

### 发明内容

[0003] 针对现有技术中的缺陷,本发明的目的是提供一种虚拟机动态保存快照的方法,其在不影响虚拟机内应用服务的前提下完成虚拟机快照操作,并能在需要时根据快照保存的档案完全恢复虚拟机磁盘与内存的当时状态。

[0004] 根据本发明的一个方面,提供一种虚拟机动态保存快照的方法,其特征在于,所述虚拟机动态保存快照的方法包括以下步骤:

[0005] 步骤一,使用一特定结构初始化文件,该文件在后续步骤用于分段保存虚拟机状态数据与内存数据,此步骤需要初始化该文件的标准I/O流对象,并注册基础读写操作集;

[0006] 步骤二,创建一块新的虚拟磁盘文件,在虚拟机主进程中注册其结构化数据,并将其置于虚拟磁盘链表的末端;

[0007] 步骤三,遍历该虚拟机中挂载的所有虚拟设备,使用不同虚拟设备注册的特定方法保存全量的虚拟设备状态数据与客户机内存数据至步骤一创建的文件中;在数据写入的开始、循环迭代及结束过程中均使用魔数分割数据片段,用于读取时识别分段恢复;

[0008] 步骤四,循环迭代的保存虚拟设备状态与虚拟机内存中变化的部分至指定的文件,并统计每次保存数据量与保存消耗时间;

[0009] 步骤五,每次循环迭代开始,首先统计所有待决的脏数据量,跳出迭代后即加锁虚拟机I/O线程,停止虚拟CPU线程,使动态的数据不再变化,安全完成最后一部分保存动作,完成虚拟机内存与虚拟设备数据的保存后,马上切换虚拟磁盘指针至新的虚拟磁盘,使后续的I/O流向新磁盘设备,而之前的磁盘设备移动为磁盘链表的倒数第二项,同时作为磁盘快

照节点,在需要恢复快照时,直接使用该节点生成虚拟机即可;

[0010] 步骤六,解锁虚拟机IO线程,并恢复虚拟CPU线程,至此虚拟机恢复运行态。

[0011] 优选地,所述虚拟机动态保存快照的方法涉及的数据对象分为三个部分:虚拟设备状态、虚拟机内存和虚拟机磁盘。

[0012] 优选地,所述步骤三的魔数分割数据片段包括以下步骤:将数据保存到文件的步骤分为预备、迭代和结尾三部分,文件前两字节依次保存快照动作整体魔数和快照版本信息,接下来排列预备段信息,依次写入预备段开始魔数、单个设备预备快照控制信息、单个设备预备快照数据信息;完成预备段数据后,循环的写入迭代段开始魔数、单个设备迭代快照控制信息、单个设备迭代快照数据信息;最后写入完成段开始魔数、单个设备完成快照控制信息、单个设备完成快照数据信息。

[0013] 与现有技术相比,本发明具有如下的有益效果:本发明在于保存虚拟机快照时,虚拟机内运行的应用业务无感知、不会被终止其服务,从而同时满足了虚拟化技术中数据安全与服务连贯性的需求。

## 附图说明

[0014] 通过阅读参照以下附图对非限制性实施例所作的详细描述,本发明的其它特征、目的和优点将会变得更明显:

[0015] 图1为本发明实施虚拟磁盘链表的示意图。

[0016] 图2为本发明实施虚拟磁盘链表增加一个磁盘后的示意图。

## 具体实施方式

[0017] 下面结合具体实施例对本发明进行详细说明。以下实施例将有助于本领域的技术人员进一步理解本发明,但不以任何形式限制本发明。应当指出的是,对本领域的普通技术人员来说,在不脱离本发明构思的前提下,还可以做出若干变形和改进。这些都属于本发明的保护范围。

[0018] 本发明将虚拟机涉及的数据对象分为三个部分——虚拟设备状态(例如虚拟CPU寄存器内容、驱动后端缓冲区数据)、虚拟机内存和虚拟机磁盘(以下简称为“磁盘”),动态保存快照就是在这三类数据不断变化的过程中,在保障数据一致的前提下将数据保存为档案,并能在快照恢复时读取档案还原数据,构造虚拟机运行场景;所以,动态虚拟机快照在实现过程中有两个技术要点:第一、快照点保存的虚拟设备状态、内存数据及虚拟机磁盘数据必须确保协调一致;第二、保存快照时虚拟机运行态的必须停机时间需要尽量短暂,以不影响虚拟机内应用服务为准。

[0019] 本发明虚拟机动态保存快照的方法包括以下步骤:

[0020] 步骤一,使用一特定结构初始化文件,该文件在后续步骤用于分段保存虚拟机状态数据与内存数据,此步骤需要初始化该文件的标准I/O流对象,并注册基础读写操作集;虚拟机磁盘格式应使用QCOW2(Copy On Write)磁盘格式,用以之后的步骤使用磁盘链表完成磁盘部分的快照,如图1所示,虚拟磁盘链表采用虚拟机、磁盘A、磁盘B;

[0021] 步骤二,创建一块新的虚拟磁盘文件,在虚拟机主进程中注册其结构化数据,并将其置于虚拟磁盘链表的末端,即将此磁盘数据结构中的磁盘链表前项指针指向目前虚拟机

使用的磁盘文件结构;如图2所示,在图1的基础上新增了一个磁盘C;

[0022] 步骤三,遍历该虚拟机中挂载的所有虚拟设备(包括虚拟CPU,虚拟内存,虚拟PCI设备等),使用不同虚拟设备注册的特定方法保存全量的虚拟设备状态数据与客户机内存数据至步骤一创建的文件中;在数据写入的开始、循环迭代及结束过程中均使用魔数分割数据片段,用于读取时识别分段恢复。其中具体的数据写入方法会使用步骤一初始化文件时注册的操作句柄执行。魔数分割数据片段包括以下步骤:将数据保存到文件的步骤分为预备、迭代和结尾三部分,文件前两字节依次保存快照动作整体魔数和快照版本信息,接下来排列预备段信息,依次写入预备段开始魔数、单个设备预备快照控制信息、单个设备预备快照数据信息;完成预备段数据后,循环的写入迭代段开始魔数、单个设备迭代快照控制信息、单个设备迭代快照数据信息;最后写入完成段开始魔数、单个设备完成快照控制信息、单个设备完成快照数据信息。

[0023] 步骤四,循环迭代的保存虚拟设备状态与虚拟机内存中变化的部分至指定的文件,并统计每次保存数据量与保存消耗时间;具体实现方式如下:在快照任务初始时记录初始时间值;每次保存数据的循环结束时取得现在时间值,并根据快照文件中记录的现有数据位置获得历史已写入的总体数据量;使用以下公式(1)计算单位时间内可保存数据量:

[0024] 单位时间内可保存数据量=已写入总体数据量/(现在时间值-初始时间值)……(1)

[0025] 其中具体的写入方法会使用步骤一初始化文件时注册的操作句柄执行。

[0026] 步骤五,每次循环迭代开始,首先统计所有待决的脏数据量,当满足如下公式(2)则判定跳出迭代:

[0027] 跳出迭代=单位时间内可保存数据量(第四步介绍)\*快照最大停机时间(默认500毫秒)>待决的脏数据量……………(2)

[0028] 跳出迭代后即加锁虚拟机IO线程,停止虚拟CPU线程,使动态的数据不再变化,安全完成最后一部分保存动作,完成虚拟机内存与虚拟设备数据的保存后,马上切换虚拟磁盘指针至新的虚拟磁盘,使后续的I/O流向新磁盘设备,而之前的磁盘设备移动为磁盘链表的倒数第二项,同时作为磁盘快照节点,在需要恢复快照时,直接使用该节点生成虚拟机即可。

[0029] 步骤六,解锁虚拟机IO线程,并恢复虚拟CPU线程,至此虚拟机恢复运行态;后续发起刷新消费虚拟磁盘IO队列的异步请求,以便完成之前磁盘设备中未处理完的IO请求。

[0030] 根据以上步骤,虚拟机快照保存已完成,虚拟机总体停机时间推算公式为式(3):

[0031] 虚拟机停机时间<=内存快照最大停机时间(500毫秒)+虚拟机磁盘指针切换时间(可忽略)+虚拟机I/O线程加解锁时间(约20毫秒)+虚拟CPU停止时间(约50毫秒)+虚拟CPU启动时间(约50毫秒)……………(3)

[0032] 根据式(3)推算虚拟机停机时间控制在620毫秒之内,对于虚拟机内的应用几乎没有影响(例如:远小于TCP超时时间)。

[0033] 以上对本发明的具体实施例进行了描述。需要理解的是,本发明并不局限于上述特定实施方式,本领域技术人员可以在权利要求的范围内做出各种变形或修改,这并不影响本发明的实质内容。

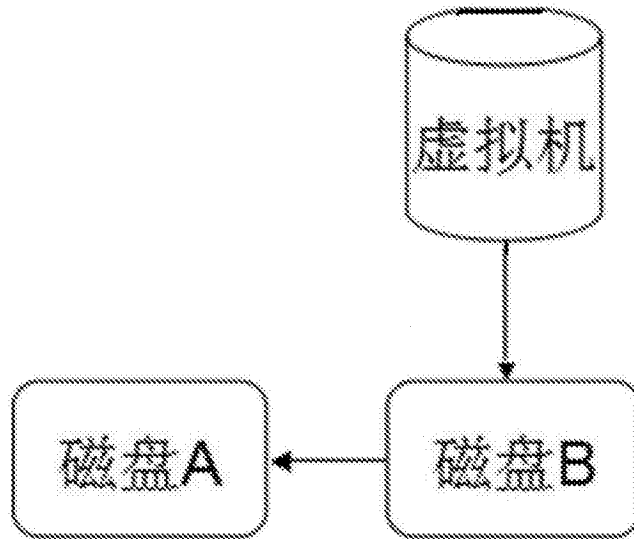


图1

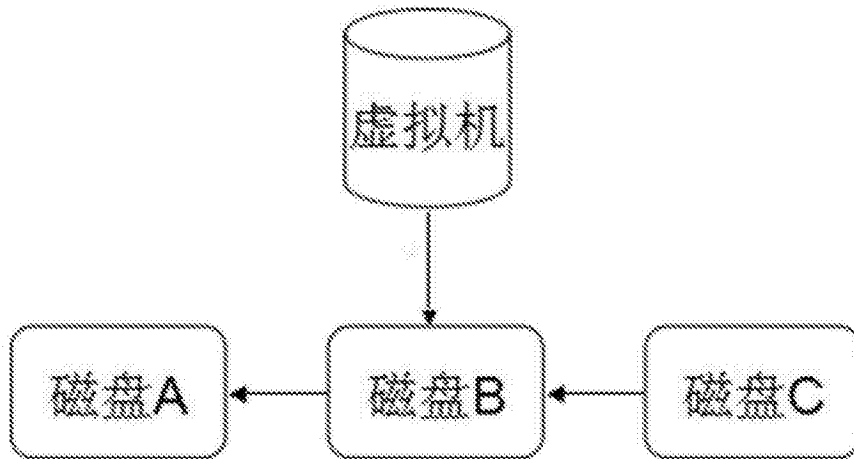


图2