

说明书

发明名称: 数据保护方法及系统、存储服务器和客户端

[0001] 本申请要求于2023年04月04日提交中国专利局、申请号为202310352365.X、申请名称为“数据保护方法及系统、存储服务器和客户端”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

[0002] 本申请涉及到云计算领域，具体而言，涉及一种数据保护方法及系统、存储服务器和客户端。

背景技术

[0003] 本部分旨在为权利要求书中陈述的本发明实施例提供背景或上下文。此处的描述不因为包括在本部分中就承认是现有技术。

[0004] 随着云计算技术的发展，越来越多的客户基于公共云构建IT(Internet Technology, 互联网技术)基础设施，虽然，主流的公共云厂商通过IaaS(Infrastructure as a Service, 基础设施即服务)层的虚拟化、网络隔离等技术可以大幅提升公共云厂商的IT资源安全性，但是，客户的虚拟机与外界公网互通，容易被病毒侵袭。为了解决这一问题，现有技术对客户的EBS(Elastic Block Storage, 弹性块存储)云盘数据进行恶意软件防护，具体过程为定期对EBS云盘打一个快照，然后将快照还原为一个新的EBS云盘，并对其进行病毒扫描。虽然该方案可以对恶意软件起到一定的事后防护作用，但是，无法实时对恶意软件进行扫描监督。同时，扫描需要消耗一定的存储成本。

[0005] 发明内容

[0006] 本申请实施例提供了一种数据保护方法及系统、存储服务器和客户端，以更加及时地识别恶意软件，降低数据安全风险。

[0007] 根据本申请的一个方面，还提供了一种数据保护方法，该方法包括：响应于接收到的读写请求，读取云盘通过目标顺序存储的数据；将所述数据转换为目标文件系统格式，按照所述目标顺序重放所述目标文件系统格式的数据；其中，所述目标文件系统格式与所述客户端的文件系统适配；扫描重放后的所述目标

文件系统格式的数据，得到恶意软件识别结果；处理所述恶意软件识别结果以恢复或保护所述云盘中存储的数据。

- [0008] 根据本申请的另一个方面，还提供了一种存储服务器，其中，包括防护模块和EBS模块；所述防护模块用于执行上述的方法步骤。
- [0009] 根据本申请的另一个方面，还提供了一种客户端，包括应用模块、文件系统和块设备；所述应用模块，用于通过所述文件系统和所述块设备向权利要求上述的存储服务器发送读写请求，以使所述EBS模块根据所述读写请求，按照目标顺序存储数据。
- [0010] 根据本申请的另一个方面，还提供了一种数据保护系统，包括客户端和存储服务器；所述存储服务器包括防护模块和EBS模块；所述防护模块用于执行上述的方法步骤；所述客户端包括应用模块、文件系统和块设备；所述应用模块，用于通过所述文件系统和所述块设备向所述存储服务器发送读写请求，以使所述EBS模块根据所述读写请求，按照目标顺序存储数据。
- [0011] 根据本申请的另一个方面，还提供了一种电子设备，包括：处理器；以及存储程序的存储器，其中，所述程序包括指令，所述指令在由所述处理器执行时使所述处理器执行根据上述的方法。
- [0012] 根据本申请的另一个方面，还提供了一种存储有计算机指令的非瞬时计算机可读存储介质，其中，所述计算机指令用于使所述计算机执行根据上述的方法步骤。
- [0013] 根据本申请的另一个方面，还提供了一种计算机程序产品，所述计算机程序产品包括计算机程序，所述计算机程序被处理器执行时实现上述的方法步骤。
- [0014] 在本申请实施例中，响应于接收到的读写请求，读取所述云盘通过目标顺序存储的数据；将所述数据转换为目标文件系统格式，按照所述目标顺序重放所述目标文件系统格式的数据；其中，所述目标文件系统格式与所述客户端的文件系统适配；扫描重放后的所述目标文件系统格式的数据，得到恶意软件识别结果；处理所述恶意软件识别结果以恢复或保护所述云盘中存储的数据。本申请响应于读写请求，读取云盘通过目标顺序存储的数据，将该数据转换成与客户端的文件系统适配的目标文件系统格式后，持续追踪扫描重放的目标文件系统

格式的数据，实现了恶意软件识别的实时性，最大限度降低数据安全风险，无需单独为扫描任务规划存储空间，降低存储成本。

附图说明

- [0015] 在下面结合附图对于示例性实施例的描述中，本公开的更多细节、特征和优点被公开，在附图中：
- [0016] 图1示出了根据本申请实施例的数据保护方法的流程图；
- [0017] 图2示出了现有云盘数据保护方案的架构图；
- [0018] 图3示出了根据本申请实施例的数据保护方法的架构图；
- [0019] 图4示出了能够用于实现本公开的实施例的示例性电子设备的结构框图；
- [0020] 图5示出了根据本申请实施例的存储服务器的结构示意图。

具体实施方式

- [0021] 下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例，然而应当理解的是，本公开可以通过各种形式来实现，而且不应该被解释为限于这里阐述的实施例，相反提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是，本公开的附图及实施例仅用于示例性作用，并非用于限制本公开的保护范围。
- [0022] 应当理解，本公开的方法实施方式中记载的各个步骤可以按照不同的顺序执行，和/或并行执行。此外，方法实施方式可以包括附加的步骤和/或省略执行示出的步骤。本公开的范围在此方面不受限制。
- [0023] 本文使用的术语“包括”及其变形是开放性包括，即“包括但不限于”。术语“基于”是“至少部分地基于”。术语“一个实施例”表示“至少一个实施例”；术语“另一实施例”表示“至少一个另外的实施例”；术语“一些实施例”表示“至少一些实施例”。其他术语的相关定义将在下文描述中给出。需要注意，本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分，并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

- [0024] 需要注意，本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的，本领域技术人员应当理解，除非在上下文另有明确指出，否则应该理解为“一个或多个”。
- [0025] 需要说明的是，本申请所涉及的用户信息(包括但不限于用户设备信息、用户个人信息等)和数据(包括但不限于用于分析的数据、存储的数据、展示的数据等)，均为经用户授权或者经过各方充分授权的信息和数据，并且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准，并提供有相应的操作入口，供用户选择授权或者拒绝。
- [0026] 本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的，而并不是用于对这些消息或信息的范围进行限制。
- [0027] 随着计算机技术的广泛应用以及互联网技术的高度普及，恶意软件也随之盛行起来。由于这些恶意软件具有高度的传染性、隐蔽性以及破坏性，已成为困扰计算机行业乃至互联网行业的一大顽疾。
- [0028] 参见图2所示的现有云盘数据保护方案的架构图，现有方案的架构中，包括：
- [0029] 1.客户虚拟机，为客户的应用程序提供计算、存储、网络等资源；
- [0030] a.应用：客户的应用程序；
- [0031] b.文件系统：客户操作系统中的文件系统，为上层应用提供数据存储服务，用于保存客户的关键数据；
- [0032] c.块设备：EBS块存储在客户虚拟机中呈现的虚拟块设备，为上层文件系统提供块接口；
- [0033] 2.远程存储服务器，作为EBS块设备的物理存储资源；
- [0034] a.EBS：作为EBS块设备的远程服务端，为计算侧的客户端提供块设备资源访问能力；
- [0035] b.存储单元：组成EBS的物理存储单元，用于将客户数据分散存储到不同的磁盘上提升数据性能以及可靠性；
- [0036] 3.快照服务，提供基于EBS块设备制作快照以及用快照还原的能力；
- [0037] 4.对象存储服务，用于保存快照文件；
- [0038] 5.临时虚拟机，用于为扫描软件提供运行环境；

- [0039] a.恶意软件扫描模块，通过开源或者商用的反病毒软件对客户的EBS数据进行安全扫描。
- [0040] 基于上述架构，按照图2中序号①-⑥的流程，执行如下步骤：
- [0041] 防护监督模块实时监督某客户EC2的异常行为，例如，与可疑的外部服务进行命令交互、被恶意攻击等；当可疑事件发生后，防护监督模块通知快照服务模块对客户的EBS制作快照；之后，快照服务模块对EBS制作快照；并将快照文件保存在远程的对象存储服务模块中；防护监督模块启动临时的EC2的虚拟机，通过快照服务模块还原一个新的EBS挂载到EC2上，并利用恶意软件扫描模块对EBS的数据进行恶意软件扫描；恶意软件扫描模块一旦发现恶意软件则通过防护监督模块上报给客户进行下一步处理。
- [0042] 该方案扫描过程需要创建快照以及一个新的EBS存储，引入额外的存储资源以及额外的计算资源开销，其中，计算资源如临时虚拟机。引入额外的资源会提升扫描成本，快照制作存在周期性，导致RPO太大，恢复数据时会造成大量的有效数据丢失，快照的恢复是整个云盘级别的，会导致有效数据大量丢失。
- [0043] 恶意软件的侵入过程十分隐蔽，很难通过预先拦截防护，另外，像勒索病毒之类的恶意软件灰度客户数据造成极大的破坏，往往在问题发现时为时已晚。虽然，当前有快照作为备份和恢复手段，但是，无法保证有效数据包含在快照中，同时，快照是基于整个磁盘的还原粒度，容易造成有效数据的大范围丢失。
- [0044] 基于以上背景，本申请提供了一种数据保护方法及系统、存储服务器和客户端，该方法可基于EBS的按照目标顺序存储的数据进行安全扫描，无需单独为扫描任务规划存储空间，降低存储成本；该方法还能基于CDP(Continuous Data Protection，持续数据保护)技术对客户数据进行恢复还原，降低RPO，减少有效数据丢失；另外，该方法还实现了文件粒度的数据恢复能力，进一步减少有效数据丢失。
- [0045] 下面，对本申请涉及的术语进行说明。
- [0046] RPO(Recovery Point Objective)：即数据恢复点目标，主要指的是服务系统所能容忍的数据丢失量。

- [0047] RTO(Recovery Time Objective): 即恢复时间目标, 主要指的是所能容忍的服务停止服务的最长时间, 也就是从灾难发生到服务系统恢复服务功能所需要的最短时间周期。
- [0048] log-structured: 一种数据存储方式, 将元数据和数据按照顺序追加写入一个环形的日志文件中。
- [0049] 在本实施例中提供了一种数据保护方法, 图1是根据本申请实施例的数据保护方法的流程图, 下面对图1中所涉及到的方法步骤进行说明。
- [0050] 步骤S102, 响应于接收到的读写请求, 读取云盘通过目标顺序存储的数据。
- [0051] 在该步骤中, 读写请求可以由客户端发出, 该客户端可以是客户虚拟机, 客户在客户虚拟机的应用上通过文件系统以及块设备向EBS下发读写请求。该读写请求可以用于按照目标顺序向云盘中存储数据, 例如, 可以根据读写请求将数据按照时间顺序存储至云盘的EBS中, 从而保证历史时刻的数据和实时的数据都被存储持久化。
- [0052] 当EBS根据读写请求完成一次数据存储后, 本申请按照EBS存储数据的目标顺序实时读取所述云盘存储的数据。其中, 目标顺序可以用于根据实际需求实现对存入云盘的数据的存入顺序的指定, 相比于按照随机顺序等方式存储, 可以防止在后续数据处理过程中, 出现因为数据相互干扰导致的数据不准确的问题。对于目标顺序具体采用的顺序, 本发明实施例对此不作具体限定。例如, 可以按照时间顺序、数据重要性顺序、标签标注的顺序等, 其中, 若是按照时间顺序, 则云盘具体可以通过log-structured方式存储数据。
- [0053] 需要说明的是, 在按照log-structured结构存储数据时, 所有的写操作都会不停地将数据添加进这个log-structured数据结构中, 而不会更新原来已有的值, 因此, 基于log-structured这样的结构特点, 可以持续并实时读取云盘通过log-structured方式存储的数据。
- [0054] 另外需要说明的是, 参见图3所示的数据保护方法的架构图, 在EBS中可以包括已经完成写入操作的持久数据和正在写入的活跃数据, 读取所述云盘通过log-structured方式存储的数据, 在具体实施时, 可以利用云盘操作接口读取EBS中的持久数据。

- [0055] 步骤S104, 将所述数据转换为目标文件系统格式, 按照所述目标顺序重放所述目标文件系统格式的数据; 其中, 所述目标文件系统格式与所述客户端的文件系统适配。
- [0056] 在该步骤中, 可以利用一个统一的文件系统抽象层, 对接客户操作系统中的各类文件, 其中, 各类文件系统, 例如ext4(Fourth extended filesystem, 第四代扩展文件系统)、xfs(一种高性能的日志文件系统)等。具体实施时, 按照客户端操作系统中的文件系统格式确定目标文件系统格式, 使目标文件系统格式与所述客户端的文件系统适配, 将从EBS读取的数据转换为该目标文件系统格式。
- [0057] 需要说明的是, 在对目标文件系统格式的数据重放处理的时候, 也需要按照上述的目标顺序进行重放, 从而保证重放后的目标文件系统格式的排列顺序, 与读取云盘中存储的数据时的读取顺序一致, 以减少后续数据处理过程中出错的概率。
- [0058] 在该步骤中, 目标文件系统格式可以是一些通用的文件系统格式, 例如, 可以为ext4、xfs和NTFS(New Technology File System, Windows NT环境的文件系统)等, 因此, 在一种可能的实施方式中, 将所述数据转换为目标文件系统格式, 可以按照如下步骤实现: 将所述数据转换为ext4文件、xfs文件或NTFS文件。之后, 再按照读取EBS中数据的目标顺序, 重放目标文件系统格式的数据。
- [0059] 需要说明的是, 由于转换成目标文件系统格式后的数据, 存储顺序可能会产生变动, 因此, 通过重放操作, 使目标文件系统格式的数据的存储顺序和读取云盘存储的数据时的顺序一致, 以便于后续对恶意软件进行识别。
- [0060] 步骤S106, 扫描重放后的所述目标文件系统格式的数据, 得到恶意软件识别结果。
- [0061] 在该步骤中, 持续追踪扫描重放后的目标文件系统格式的数据, 由于不同的恶意软件存在不同的行为特征, 例如, 大量删除数据、大量读取数据、大量破坏数据以及大量修改数据等等, 根据目标文件系统格式的数据中, 针对数据的操作行为, 识别是否存在恶意软件, 得到恶意软件识别结果。
- [0062] 需要说明的是, 恶意软件是一个概括性的定义, 按照其造成破坏的性质可以分成两类: 第一类主要包括后门程序、键盘记录器、密码盗取、间谍软件等, 寄

宿在客户计算机中，通过盗取网银密码、隐私照片等重要数据来危害客户利益；第二类是一种被称为“勒索病毒”的新型电脑病毒，主要以邮件、程序木马等方式传播。这种病毒利用各种加密算法对文件进行不可逆的加密，被感染者一般无法解密，必须拿到病毒发布者的私钥解密。病毒发布者以客户的重要数据来勒索高额赎金。在该步骤中，通过扫描重放后的所述目标文件系统格式的数据，可以识别出上述两类恶意软件中的任一种，除上述列出的恶意软件种类，还可以识别出其它具有一定行为特征的恶意软件。

[0063] 在上述步骤中，可以实时发现恶意软件的侵入行为。另外，该步骤中，可直接在数据存储的生产区完成扫描，无需单独为扫描任务规划存储空间，降低存储成本。

[0064] 需要说明的是，恶意软件库中的信息可以不断进行更新，更新策略可以根据实际需求确定，本发明实施例在此不作具体限定。

[0065] 步骤S108，处理所述恶意软件识别结果以恢复或保护所述云盘中存储的数据。

[0066] 在该步骤中，恶意软件识别结果可以用于确定被破坏的数据，针对被破坏的数据，可以执行删除以及恢复等操作，从而恢复或保护云盘中存储的数据，降低数据的安全风险，保证数据的安全性。其中，具体采用的处理手段，可以根据实际需求确定，本发明实施例对此不做具体限定。

[0067] 在本申请实施例中，响应于接收到的读写请求，读取云盘通过目标顺序存储的数据；将所述数据转换为目标文件系统格式，按照所述目标顺序重放所述目标文件系统格式的数据；其中，所述目标文件系统格式与所述客户端的文件系统适配；扫描重放后的所述目标文件系统格式的数据，得到恶意软件识别结果；处理所述恶意软件识别结果以恢复或保护所述云盘中存储的数据。本申请响应于读写请求，读取云盘按照目标顺序存储的数据，将该数据转换成与客户端的文件系统适配的目标文件系统格式后，持续追踪扫描重放的目标文件系统格式的数据，实现了恶意软件识别的实时性，最大限度降低数据安全风险，无需单独为扫描任务规划存储空间，降低存储成本。

[0068] 在一种可能的实施方式中，扫描重放后的所述目标文件系统格式的数据，得到恶意软件识别结果，可以按照如下步骤执行：扫描重放后的所述目标文件系统

格式的数据，得到扫描信息；利用恶意软件库识别所述扫描信息中的目标恶意软件和被所述目标恶意软件破坏的目标文件；所述恶意软件库用于提供所述目标恶意软件的行为特征信息；将所述目标恶意软件和/或所述目标文件作为恶意软件识别结果。

[0069] 在该可能的实施方式中，扫描重放后的所述目标文件系统格式的数据得到扫描信息。可以预先生成恶意软件库，该恶意软件库用于提供目标恶意软件的文件以及行为特征信息，作为判断上述扫描信息中恶意软件以及被感染文件的依据，即利用恶意软件库识别所述扫描信息中的目标恶意软件和被所述目标恶意软件破坏的目标文件，将所述目标恶意软件和/或所述目标文件作为恶意软件识别结果。

[0070] 在识别到恶意软件后，为了防止恶意软件的破坏程度进一步加强或破坏数据量进一步增加，在一种可能的实施方式中，处理所述恶意软件识别结果，可以按照如下步骤执行：根据所述恶意软件识别结果生成处理指令，将所述处理指令发送至控制端；其中，所述处理指令用于删除和/或隔离处理所述恶意软件识别结果。

[0071] 在本发明实施例中，根据所述恶意软件识别结果，可以定位到识别出的恶意软件或被恶意软件破坏的文件，生成处理指令，并将该指令发送至控制端，以使控制端根据该处理指令删除和/或隔离处理识别出的恶意软件或被恶意软件破坏的文件。需要说明的是，控制端可以是用户端的一部分，通过用户端进行控制，也可以是一个独立的控制终端，可以根据实际需求进行设置，本发明实施例在此不作具体限定。

[0072] 如果发现被恶意软件破坏的文件，且文件中包含重要数据，为了降低有效数据的损失，提供更高的RPO，在一种可能的实施方式中，处理所述恶意软件识别结果，可以按照如下步骤执行：确定被侵入文件的恢复时刻，按照所述恢复时刻恢复所述被侵入文件或删除所述被侵入文件。

[0073] 在本发明实施例中，可以结合CDP技术，并根据实际需求确定被侵入文件的恢复时刻，具体恢复的时间点取决于不同策略，例如，可以恢复到文件刚刚创建的时刻、恢复到恶意软件侵入的时刻前一段时间、恢复到客户自定义的任意时

刻或直接清楚文件，之后，按照所述恢复时刻恢复所述被侵入文件或删除所述被侵入文件。

- [0074] 需要说明的是，CDP为用户提供了新的数据保护手段，系统管理者无须关注数据的备份过程(因为CDP系统会不断监测关键数据的变化，从而不断地自动实现数据的保护)，而且仅仅当灾难发生后，简单地选择需要恢复到的时间点即可实现数据的快速恢复。
- [0075] CDP技术通过在操作系统核心层中植入文件过滤驱动程序，来实时捕获所有文件访问操作。对于需要CDP连续备份保护的文件，当CDP管理模块经由文件过滤驱动拦截到其改写操作时，则预先将文件数据变化部分连同当前的系统时间戳(System Time Stamp)一起自动备份到CDP存储体。从理论上说，任何一次的文件数据变化都会被自动记录，因而称之为持续数据保护。
- [0076] 在上述步骤中，可以实现对有效数据更加精确的保护，实现了文件粒度的数据恢复能力，进一步减少有效数据丢失。
- [0077] 如果恶意软件侵入后，并没有重要数据需要保护，在一种可能的实施方式中，处理所述恶意软件识别结果，则可以按照如下步骤执行：若所述恶意软件识别结果中的任意文件均未达到保存条件，则恢复所述云盘的目标存储区至未被所述恶意软件侵入的状态。
- [0078] 在本发明实施例中，保存条件可以根据实际需求进行设置，用于筛选出需要保存的文件。目标存储区可以是云盘的整个磁盘。恶意软件侵入后，云盘中若不存在重要数据，即恶意软件识别结果中的任意文件均未达到保存条件，则恢复所述云盘的目标存储区至未被所述恶意软件侵入的状态，例如，恢复到恶意软件侵入的时刻前一段时间。
- [0079] 需要说明的是，当恶意软件识别结果中的任意文件均未达到保存条件时，可以启动磁盘粒度的修复功能，可以直接将整个磁盘恢复到未被侵入的状态。若恶意软件识别结果中任一文件达到保存条件，则可以按照前述的文件粒度的修复功能进行恢复操作。
- [0080] 考虑到恶意软件可能存在例如大量删除、读取或者破坏等具有一定特征的行为，为了能对潜在的侵入行为进行拦截，该方法还可以执行如下步骤：

- [0081] 利用机器学习模型和恶意行为特征库，基于所述通过目标顺序存储的数据，确定风险行为；生成拦截指令，将所述拦截指令发送至控制端，以使所述控制端拦截所述风险行为。
- [0082] 在本发明实施例中，恶意行为特征库用于提供风险行为的特征信息，可以利用恶意行为特征库预先训练机器学习模型。利用训练好的机器学习模型实时识别出通过目标顺序存储的数据中的异常行为，即利用机器学习模型和恶意行为特征库，基于所述通过目标顺序存储的数据，确定风险行为。如果发现异常则触发对潜在的侵入行为的拦截操作，即生成拦截指令，将所述拦截指令发送至控制端，以使所述控制端拦截所述风险行为。
- [0083] 需要说明的是，控制端可以是用户端的一部分，通过用户端进行控制，也可以是一个独立的控制终端，可以根据实际需求进行设置，本发明实施例在此不作具体限定。
- [0084] 本申请还提供了一种存储服务器，参见图5所示的存储服务器的结构示意图，该存储服务器500包括防护模块501和EBS模块502；所述防护模块用于执行上述的任一种数据保护方法的步骤。数据保护方法的具体实施方式在此不再赘述。
- [0085] 参见图3所示的数据保护方法的架构图，下面对存储服务器的一种可能的架构及该架构的工作流程进行说明。
- [0086] 存储服务器中上述的防护模块可以包括如下架构：
- [0087] a.恶意软件扫描模块，负责对客户云盘数据进行恶意软件扫描，并精确识别出恶意软件以及被恶意软件破坏的文件；
- [0088] i.恶意软件库，提供恶意软件的文件以及行为特征，作为判断恶意软件以及被感染文件的依据；
- [0089] ii.文件系统适配装置(图3中的通用文件接口层)，用于提供一个统一的文件系统抽象层，对接客户操作系统中的各类文件系统；
- [0090] iii.云盘操作接口，用于访问云盘数据；
- [0091] b.数据恢复装置：当用户数据被恶意软件侵入后，负责将数据还原到被侵入之前的状态，或者清除被破坏的文件；

- [0092] i.文件恢复模块，用于文件粒度的修复功能，可以精确的将恶意文件以及被感染文件清除或者直接还原到未被感染前的状态；
- [0093] ii.整盘恢复模块，用于磁盘粒度的修复功能，可以直接将整个磁盘恢复到未被侵入的状态；
- [0094] iii.持续数据保护模块：用于提供数据的持续保护能力，提供任意时刻的数据还原功能；
- [0095] c.恶意行为预测装置：基于恶意软件特征库以及机器学习算法，对恶意软件的数据破坏行为进行提前预判；
- [0096] i.恶意行为特征库：负责跟踪客户数据的操作行为，预判可能出现的恶意软件破坏行为；
- [0097] ii.数据保护触发模块：当发现潜在的数据破坏风险后，负责通知Scanner对潜在的侵入行为进行拦截；
- [0098] 存储服务器中上述的EBS模块，可以包括如下架构：
- [0099] a.持久数据：基于log-structured数据结构，记录的客户云盘历史数据；
- [0100] b.活跃数据：客户云盘写入的最新数据。
- [0101] 基于图3展示的架构，按照图3中①—③的流程，通过如下步骤，可以实现上述的数据保护方法：
- [0102] 1.客户在客户虚拟机的应用通过文件系统以及块设备层，向EBS下发块设备读写请求；
- [0103] 2.EBS将数据通过log-structured方式存储，保证历史时刻的数据和实时的数据都被存储持久化；
- [0104] 3.扫描装置按照时间顺序，通过云盘操作接口读取持久数据；
- [0105] 4.扫描装置通过通用文件接口层将EBS的log-structured数据格式转成通用文件系统格式，例如，ext4、xfs等，并按时间顺序重放客户数据；
- [0106] 5.扫描装置通过恶意软件扫描模块对重放后的数据进行安全扫描，实时发现恶意软件的侵入行为，并通知数据恢复装置对数据进行恢复或者隔离：
- [0107] a.如果发现恶意软件则通知客户对其进行删除或者隔离处理；

- [0108] b.如果发现被恶意软件破坏的文件，则通过文件恢复模块对文件进行恢复，具体恢复的时间点取决于不同策略，例如，1)可以恢复到文件刚刚创建的时刻，2)恢复到恶意软件侵入的时刻前一段时间，3)直接清除文件等，4)或者恢复到客户自定义的任任意时刻；
- [0109] c.如果恶意软件侵入后，并没有重要数据需要保护，则可以通过整盘恢复模块对整个磁盘进行恢复，例如，恢复到恶意软件侵入的时刻前一段时间；
- [0110] 6.同时，恶意行为预测装置基于客户的数据特征实时检测异常行为，如果发现异常则触发扫描装置对潜在的侵入行为进行拦截。
- [0111] 在利用该存储服务器，实现数据保护方法的过程中，通过对log structured数据进行持续追踪扫描，实现了恶意软件识别的实时性，最大限度减少数据安全风险，另外，直接在数据存储的生产区完成扫描，无需单独为扫描任务规划存储空间，降低存储成本，提供更低的RPO(Recovery Point Objective)和RTO(Recovery Time Objective)能力，当数据被恶意软件侵入后，最大限度降低有效数据的损失。本申请在存储成本上有大幅降低，同时，解决了扫描任务对客户环境的侵入问题，另外，可以更加及时的识别恶意软件，并对有效数据进行更加精确的保护。本申请基于公共云块存储的log structured存储结构特点，实现了智能化的恶意软件防护，可以应用在涉及云计算的EBS存储产品中，以及云安全相关产品中。
- [0112] 本申请还提供了一种客户端，包括应用模块、文件系统和块设备；所述应用模块，用于通过所述文件系统和所述块设备向上述的存储服务器发送读写请求，以使所述EBS模块根据所述读写请求，按照目标顺序存储数据。
- [0113] 需要说明的是，该客户端可以是客户虚拟机，为客户的应用程序提供计算、存储、网络等资源。其中，文件系统可以是客户操作系统中的文件系统，为上层应用提供数据存储服务，用于保存客户的关键数据。块设备可以是EBS块存储在客户虚拟机中呈现的虚拟块设备，为上层文件系统提供块接口。存储服务器，可以作为EBS块设备的物理存储资源。
- [0114] 本申请还提供了一种数据保护系统，包括客户端和存储服务器；所述存储服务器包括防护模块和EBS模块；所述防护模块用于执行上述的方法步骤；所述客户

端包括应用模块、文件系统和块设备；所述应用模块，用于通过所述文件系统和所述块设备向所述存储服务器发送读写请求，以使所述EBS模块根据所述读写请求，按照目标顺序存储数据。

[0115] 本公开还提供了一种数据保护装置，该装置包括：接口模块，用于响应于接收到的读写请求，读取云盘通过目标顺序存储的数据；

[0116] 文件转换模块，用于将所述数据转换为目标文件系统格式，按照所述目标顺序重放所述目标文件系统格式的数据；其中，所述目标文件系统格式与所述客户端的文件系统适配；恶意文件扫描模块，用于扫描重放后的所述目标文件系统格式的数据，得到恶意软件识别结果；数据恢复模块，用于处理所述恶意软件识别结果以恢复或保护所述云盘中存储的数据。

[0117] 该系统或者装置用于实现上述的实施例中的方法的功能，该系统或者装置中的每个模块与方法中的每个步骤相对应，已经在方法中进行过说明的，在此不再赘述。

[0118] 可选地，将所述数据转换为目标文件系统格式，包括：将所述数据转换为ext4文件、xfs文件或NTFS文件。

[0119] 可选地，扫描重放后的所述目标文件系统格式的数据，得到恶意软件识别结果，包括：扫描重放后的所述目标文件系统格式的数据，得到扫描信息；利用恶意软件库识别所述扫描信息中的目标恶意软件和被所述目标恶意软件破坏的目标文件；所述恶意软件库用于提供所述目标恶意软件的行为特征信息；将所述目标恶意软件和/或所述目标文件作为恶意软件识别结果。

[0120] 可选地，处理所述恶意软件识别结果，包括：根据所述恶意软件识别结果生成处理指令，将所述处理指令发送至控制端；其中，所述处理指令用于删除和/或隔离处理所述恶意软件识别结果。

[0121] 可选地，处理所述恶意软件识别结果，包括：确定被侵入文件的恢复时刻，按照所述恢复时刻恢复所述被侵入文件或删除所述被侵入文件。

[0122] 可选地，处理所述恶意软件识别结果，包括：若所述恶意软件识别结果中的任意文件均未达到保存条件，则恢复所述云盘的目标存储区至未被所述恶意软件侵入的状态。

- [0123] 可选地，还包括：利用机器学习模型和恶意行为特征库，基于所述通过目标顺序存储的数据，确定风险行为；生成拦截指令，将所述拦截指令发送至控制端，以使所述控制端拦截所述风险行为。
- [0124] 本公开示例性实施例还提供一种电子设备，包括：至少一个处理器；以及与至少一个处理器通信连接的存储器。所述存储器存储有能够被所述至少一个处理器执行的计算机程序，所述计算机程序在被所述至少一个处理器执行时用于使所述电子设备执行根据本公开实施例的方法。
- [0125] 本公开示例性实施例还提供一种存储有计算机程序的非瞬时计算机可读存储介质，其中，所述计算机程序在被计算机的处理器执行时用于使所述计算机执行根据本公开实施例的方法。
- [0126] 本公开示例性实施例还提供一种计算机程序产品，包括计算机程序，其中，所述计算机程序在被计算机的处理器执行时用于使所述计算机执行根据本公开实施例的方法。
- [0127] 参考图4，现将描述可以作为本公开的服务器或客户端的电子设备400的结构框图，其是可以应用于本公开的各方面的硬件设备的示例。电子设备旨在表示各种形式的数字电子的计算机设备，诸如，膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置，诸如，个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例，并且不意在限制本文中描述的和/或者要求的本公开的实现。
- [0128] 如图4所示，电子设备400包括计算单元401，其可以根据存储在只读存储器(ROM)402中的计算机程序或者从存储单元408加载到随机访问存储器(RAM)403中的计算机程序，来执行各种适当的动作和处理。在RAM 403中，还可存储设备400操作所需的各种程序和数据。计算单元401、ROM 402以及RAM 403通过总线404彼此相连。输入/输出(I/O)接口405也连接至总线404。
- [0129] 电子设备400中的多个部件连接至I/O接口405，包括：输入单元406、输出单元407、存储单元408以及通信单元409。输入单元406可以是能向电子设备400输入

信息的任何类型的设备，输入单元406可以接收输入的数字或字符信息，以及产生与电子设备的用户设置和/或功能控制有关的键信号输入。输出单元407可以是能呈现信息的任何类型的设备，并且可以包括但不限于显示器、扬声器、视频/音频输出终端、振动器和/或打印机。存储单元408可以包括但不限于磁盘、光盘。通信单元409允许电子设备400通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据，并且可以包括但不限于调制解调器、网卡、红外通信设备、无线通信收发机和/或芯片组，例如蓝牙设备、WiFi设备、WiMax设备、蜂窝通信设备和/或类似物。

- [0130] 计算单元401可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元401的一些示例包括但不限于处理单元(CPU)、图形处理单元(GPU)、各种专用的人工智能(AI)计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器(DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元401执行上文所描述的各个方法和处理。例如，在一些实施例中，前述数据保护方法可被实现为计算机软件程序，其被有形地包含于机器可读介质，例如存储单元408。在一些实施例中，计算机程序的部分或者全部可以经由ROM 402和/或通信单元409而被载入和/或安装到电子设备400上。在一些实施例中，计算单元401可以通过其他任何适当的方式(例如，借助于固件)而被配置为执行数据保护方法。
- [0131] 用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器或控制器，使得程序代码当由处理器或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行，作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。
- [0132] 在本公开的上下文中，机器可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备，或者上述内容的任何合适组合。机器可读存储介质的更

具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM或快闪存储器)、光纤、便捷式紧凑盘只读存储器(CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

- [0133] 如本公开使用的，术语“机器可读介质”和“计算机可读介质”指的是用于将机器指令和/或数据提供给可编程处理器的任何计算机程序产品、设备、和/或装置(例如，磁盘、光盘、存储器、可编程逻辑装置(PLD))，包括，接收作为机器可读信号的机器指令的机器可读介质。术语“机器可读信号”指的是用于将机器指令和/或数据提供给可编程处理器的任何信号。
- [0134] 为了提供与用户的交互，可以在计算机上实施此处描述的系统和技术，该计算机具有：用于向用户显示信息的显示装置(例如，CRT(阴极射线管)或者LCD(液晶显示器))；以及键盘和指向装置(例如，鼠标或者轨迹球)，用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互；例如，提供给用户的反馈可以是任何形式的传感反馈(例如，视觉反馈、听觉反馈、或者触觉反馈)；并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。
- [0135] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如，作为数据服务器)、或者包括中间件部件的计算系统(例如，应用服务器)、或者包括前端部件的计算系统(例如，具有图形用户界面或者网络浏览器的用户计算机，用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如，通信网络)来将系统的部件相互连接。通信网络的示例包括：局域网(LAN)、广域网(WAN)和互联网。
- [0136] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务器的关系。

[0137] 以上仅为本申请的实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本申请的权利要求范围之内。

权利要求书

- [权利要求 1] 一种数据保护方法，包括：
- 响应于接收到的读写请求，读取云盘通过目标顺序存储的数据；
- 将所述数据转换为目标文件系统格式，按照所述目标顺序重放所述目标文件系统格式的数据；其中，所述目标文件系统格式与所述客户端的文件系统适配；
- 扫描重放后的所述目标文件系统格式的数据，得到恶意软件识别结果；
- 处理所述恶意软件识别结果以恢复或保护所述云盘中存储的数据。
- [权利要求 2] 根据权利要求1所述的方法，其中，将所述数据转换为目标文件系统格式，包括：
- 将所述数据转换为ext4文件、xfs文件或NTFS文件。
- [权利要求 3] 根据权利要求1所述的方法，其中，扫描重放后的所述目标文件系统格式的数据，得到恶意软件识别结果，包括：
- 扫描重放后的所述目标文件系统格式的数据，得到扫描信息；
- 利用恶意软件库识别所述扫描信息中的目标恶意软件和被所述目标恶意软件破坏的目标文件；所述恶意软件库用于提供所述目标恶意软件的行为特征信息；
- 将所述目标恶意软件和/或所述目标文件作为恶意软件识别结果。
- [权利要求 4] 根据权利要求1所述的方法，其中，处理所述恶意软件识别结果，包括：
- 根据所述恶意软件识别结果生成处理指令，将所述处理指令发送至控制端；其中，所述处理指令用于删除和/或隔离处理所述恶意软件识别结果。
- [权利要求 5] 根据权利要求1所述的方法，其中，处理所述恶意软件识别结果，包括：
- 确定被侵入文件的恢复时刻，按照所述恢复时刻恢复所述被侵入文件或删除所述被侵入文件。

- [权利要求 6] 根据权利要求1所述的方法，其中，处理所述恶意软件识别结果，包括：
若所述恶意软件识别结果中的任意文件均未达到保存条件，则恢复所述云盘的目标存储区至未被所述恶意软件侵入的状态。
- [权利要求 7] 根据权利要求1-6任一项所述的方法，其中，所述方法还包括：
利用机器学习模型和恶意行为特征库，基于所述通过目标顺序存储的数据，确定风险行为；
生成拦截指令，将所述拦截指令发送至控制端，以使所述控制端拦截所述风险行为。
- [权利要求 8] 一种存储服务器，其中，包括防护模块和EBS模块；所述防护模块用于执行权利要求1-7任一项所述的方法步骤。
- [权利要求 9] 一种客户端，包括应用模块、文件系统和块设备；
所述应用模块，用于通过所述文件系统和所述块设备向权利要求8所述的存储服务器发送读写请求，以使所述EBS模块根据所述读写请求，按照目标顺序存储数据。
- [权利要求 10] 一种数据保护系统，包括客户端和存储服务器；
所述存储服务器包括防护模块和EBS模块；所述防护模块用于执行权利要求1-7任一项所述的方法步骤；
所述客户端包括应用模块、文件系统和块设备；所述应用模块，用于通过所述文件系统和所述块设备向所述存储服务器发送读写请求，以使所述EBS模块根据所述读写请求，按照目标顺序存储数据。
- [权利要求 11] 一种电子设备，包括：
处理器；以及
存储程序的存储器，
其中，所述程序包括指令，所述指令在由所述处理器执行时使所述处理器执行根据权利要求1-7中任一项所述的方法步骤。

- [权利要求 12] 一种存储有计算机指令的非瞬时计算机可读存储介质，其中，所述计算机指令用于使所述计算机执行根据权利要求1-7中任一项所述的方法步骤。
- [权利要求 13] 一种计算机程序产品，其特征在于，所述计算机程序产品包括计算机程序，所述计算机程序被处理器执行时实现权利要求1-7中任一项所述的方法步骤。