

云计算环境下的取证问题研究

何晓行¹ 王剑虹²

(重庆邮电大学法学院 重庆 400065)¹ (西南政法大学法学院 重庆 401120)²

摘 要 云计算是一种基于互联网的超级计算模式,它的快速发展给取证带来了一定的挑战。在云计算环境下,取证活动包括证据发现、证据固定及证据提取 3 个阶段。但云计算自身存在的安全风险,导致云计算环境下的取证活动亦面临一定的技术及法律问题。对于取证信息及手段方面的问题,可以采取数据迁移技术及入侵检测系统等手段加以解决。对于相关的法律问题,则可以采取完善相关法律法规、强化国际区际合作等途径加以解决。

关键词 云计算,取证,数据迁移,虚拟化

Research on Evidence Collection under Cloud Computing Environment

HE Xiao-xing¹ WANG Jian-hong²

(Law Institute, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)¹

(Law Institute, Southwest University of Political Science and Law, Chongqing 401120, China)²

Abstract Cloud computing is a kind of super computing model on the internet. The development of cloud computing brings some challenges to evidence collection. Under cloud computing environment, the evidence collection includes three stages: detection, fixup and scavenge. There are some safe risks in cloud computing, so collecting evidence under cloud computing environment will face some technologic and legal problems. As to the problems on evidence information and collecting measure, we can solve them by the technology of data migration and Intrusion Detection System. As to the legal problems, we can solve them by improving the legislation and strengthening the international cooperation.

Keywords Cloud computing, Collecting evidence, Data migration, Virtualization

随着计算机软硬件技术的不断发展,各种新的概念和架构也得到不断的发展,而云计算技术即为近年来计算机领域提出的最富革命意义的概念之一。目前,云计算技术已经逐步发展成熟,并在很多重要的应用领域取得了成功。将云技术应用于计算机取证领域,则是计算机取证模式发展的一个重要趋势。但是云计算的快速发展,在带来了巨大经济收益的同时,也对电子取证专业人员带了巨大的挑战。本文将围绕云计算环境下的取证问题,分析云计算环境下取证面临的技术及法律问题,并提出相应的对策。

1 云计算的概念与特点

1.1 云计算的概念

“云计算”一词由 Google 的创始人谢尔盖·布林和拉里·佩奇最早提出,它是继服务计算和网格计算之后的一种新兴的商业计算模型,也可以说是网格计算(Grid Computing)、分布式计算(Distributed Computing)、并行计算(Parallel Computing)、效用计算(Utility Computing)、网络存储(Network Storage Technologies)、虚拟化(Virtualization)、负载均衡(Load Balance)等计算机技术和网络技术的融合与发展。对于“云计算”概念的表述,不同的人员及组织存在一定的区别。

我国网格计算、云计算专家刘鹏将“云计算”的概念界定为:“云计算将计算任务分布在大量计算机构成的资源池上,使用户能够按需获取计算力、存储空间和信息服务”^[1]。而 IBM 公司在其云计算计划的技术白皮书中对“云计算”的概念作出如下描述,即“云计算一词用来同时描述一个系统平台或者一种类型的应用程序。一个云计算的平台按需进行动态部署(provision)、配置(configuration)、重新配置(reconfigure)以及取消服务(deprovision)等”^[2]。另外,云计算的一个核心理念就是通过不断提高“云”的处理能力,进而减少用户终端的处理负担,最终使用户终端简化成一个单纯的输入输出设备,并能按需享受“云”的强大计算处理能力。著名咨询机构 Gartner 将云计算定义为“云计算是利用互联网技术将庞大且可伸缩的 IT 能力集合起来作为服务提供给多个客户的技术”^[3]。

综上所述,云计算的核心是海量数据的存储和计算,由几十万台甚至几百万台计算机构成的计算机群对信息进行聚合和分布处理,然后通过网络对客户提供服务^[4]。云计算旨在通过网络把多个成本相对较低的计算实体整合成一个具有强大计算能力的完美系统,并借助 SaaS、PaaS、IaaS、MSP 等先进的商业模式把这强大的计算能力分布到终端用户手中。它

到稿日期:2011-12-20 返修日期:2012-03-10 本文受西南政法大学校级重点项目(2010-XZZD25),重庆市社会科学规划青年项目(2011QNFX52)资助。

何晓行(1975—),女,博士,讲师,主要研究方向为电子证据与计算机应用技术,E-mail:morninghe@163.com;王剑虹(1978—),女,博士,副教授,主要研究方向为电子证据与计算机应用技术。

通过网络将庞大的计算处理程序自动分拆成无数个较小的子程序,再交由多部服务器所组成的庞大系统,经搜寻、计算、分析之后将处理结果回传给用户。使用这项技术,网络服务提供者可以在数秒之内处理数以千万计甚至亿计的信息,实现与超级计算机同样强大效能的网络服务。

1.2 云计算的特点

与传统的单机或网络应用模式相比,云计算具有如下显著特点。

1.2.1 数据的安全可靠性

云计算提供了最可靠、最安全的数据存储中心,用户无需担心数据丢失、病毒入侵等麻烦。“云”使用了数据多副本容错、计算节点同构可互换等措施来保障服务的高可靠性,使用云计算比使用本地计算机更可靠。具体而言,当文档、资料被保存在“云”网上时,在“云”的另一端,有先进的数据存储中心和专业的团队来管理信息,同时,严格的权限管理策略可以进一步促进数据共享。

1.2.2 规模的高扩展性

高扩展性表达了云计算能够无缝地扩展到大规模的集群之上,甚至包含数千个节点同时进行处理^[5]。提供计算资源的“云”的规模可以根据实际应用中的需求动态伸缩,以此满足应用和用户的弹性需求。这一特点决定了云计算环境中的实时取证模型也应当具有可扩展性,其可根据实际取证需要动态伸缩,满足安全人员的取证需要^[6]。

1.2.3 数据存储的分布性

在云计算环境中,客户使用的资源均不在本地,其进行计算时所涉及的数据信息会分布于“云”中各个服务器中,故其所使用的数据之存储具有分布性,这一特点也决定了云计算环境中的取证模型在监测各个公共云使用状况的同时,对可疑云资源中数据存储的确切位置要有实时的记录和分析功能。

1.2.4 系统的虚拟化

系统虚拟化技术是云计算平台最核心的技术,云计算平台的产生就是系统虚拟化技术发展的产物^[7]。系统虚拟化主要是提供一个虚拟化的环境,在这个虚拟化的环境中,所有的执行环境都使用同一个硬件平台,每一个执行环境都认为自己独占硬件平台运行。为了保证每个执行环境的相互独立性,首先进行软件和硬件的划分,利用仿真技术,在时间上分隔运行,同时保证执行环境的应用服务能力,这样在同一个硬件平台上可以同时运行多个相互隔离的系统执行环境^[8]。云计算支持用户在任意位置使用各种终端访问云计算平台,用户所请求的资源不是来自某一特定的有形实体,而是在提供资源的“云”中获得。在云服务部署层,提供商通过用户访问接口来对各个用户进行有效的资源分配和访问控制,并对云计算平台的虚拟机资源进行统一管理和调度;虚拟化层的作用主要是利用系统虚拟化技术,将各种分散的物理计算资源进行虚拟整合,提供给服务部署层。

从取证角度来看,云计算平台是由大量分布式异构虚拟计算资源构建的平台,在整合计算资源且保持低成本的状态下提供良好的计算服务质量,使企业和个人用户都能通过云计算的海量信息库来实现信息的自由分享,这在一定程度上给计算机取证工作带来较大的挑战。

2 云计算环境下的取证实现

在传统的取证模式下,取证主要是针对具体的设备和文件对取证对象进行隔离和保护,整个取证过程基本上在一个静止的封闭空间中进行,取证对象明确,没有其他数据干扰。而在云计算环境下,由于其主要由大量的分布式异构虚拟计算资源构成,而且多个用户共用数据和资源,这些用户可能协同工作,也可能相互提供服务,产生各种临时数据和文件访问记录,因此在云计算环境下,电子证据取证的实现与传统的取证模式具有显著区别。在云计算环境下,取证的实现主要包括证据发现、证据固定及证据提取3个阶段。

2.1 证据发现

证据发现为计算机取证的准备阶段,即通过侦查和现场搜集获得原始的数据信息。在云计算环境下,计算机的使用者利用手中的终端设备相互联网上的“云”服务器发送指令信息,在网络上使用云服务。在这一过程中,几乎全部操作都要在网络的节点中经由特定的第三方“智能程序”按照指令自动完成,并且依据严格、统一的“协议”运行并踏实地记录下第一个信息处理。因此,使用者的整个操作就反映为两部分内容:一是自己主观上想要编辑、浏览和运算的数据信息;二是由一系列客观的“智能程序”自动、严格、准确地记录下的信息流动记录——操作痕迹^[9]。因此,在云计算环境下,可以从以下几个途径来发现证据。

首先,可以从云端调查着手发现相关证据。在云计算环境下,用户的很多行为都是直接在网上完成,如Google的私有云提供了在线图片的存储和编辑、文档的编辑、表格的编辑,因此操作者的相关操作痕迹必然会在信息管理数据库中留下较为详细的记录,从而“云”操作的每一个步骤就得以详细记录,如时间、帐户信息、存储信息等。由此可见,取证者在获得司法授权的前提下,可以直接从云计算开发公司获取节点数据,从该节点数据处去查找网络环境下的涉案使用痕迹和违法证据。

其次,从客户端着手发现相关证据。虽然在云计算下很多数据都是在网络上传递,但是最终这些数据还得为用户服务,因此客户端主要会存在很多的动态信息,比如内存信息、进程状况、缓冲区信息、文件碎片等。若犯罪嫌疑人云中从事犯罪活动,其本地客户端势必会遗留部分残余电子数据^[10],可以从中发现、查找相关证据。

再次,从政府监管角度着手发现证据。不管是私有云还是公有云,也不管服务提供商是国内还是国外,都必须符合最终落地国的法律法规。因此我们也可以利用服务提供商在所属地审计控制节点处进行取证调查,比如,目前国内研究云计算最早的中科院高能物理研究所与欧洲及一些大学合作的项目,它的接入点就在上海等三地;而且根据不同文化过滤了不同的词汇,我们可以直接从这个接入点去查找涉案的信息,然后根据这些信息去追踪相关的IP地址^[11]。

最后,从提供云下服务的子云发现证据。云服务提供商在云计算开发方提供的云下提供子云(子服务)。数据信息会由服务提供商所记录,这些信息记录可能会包含操作者的IP地址、操作时间、使用系统及终端型号等内容,这对于查明某些案件的基本事实具有关键作用。这在某些案件的重要性和

数量达不到去云计算开发方那里取证调查时,可以从云服务提供商着手发现证据。云服务提供商由于在地域范围方面相对较小,因此能在一定程度上降低取证成本。

2.2 证据固定

证据固定主要是解决证据的完整性验证,即在计算机取证的同时,通过数字签名和见证人签名等方式保证收集的证据的完整性和真实性,使证据的获取符合取证规范的要求。由于电子证据自身的特点,对其进行固定时,一方面要妥善保存电子数据信息本身的内容,另一方面要重视对其周边环境数据的保全,尤其是那些能够证明电子数据提取时各种存在环境的信息。如可以采取迁移取证监管来调度和监控镜像证据提取层的每一次迁移操作,并记录下全部的迁移过程信息,保证迁移过程中的策略选择和执行控制,并生成监管过程报告,保证取证数据的可靠性和法律完整性^[12]。

2.3 证据提取

证据提取是指专业取证人员通过各种取证工具,利用数据恢复技术和解码解密技术分析原始数据,还原相关的隐藏信息,恢复被删除的有价值的文件。如果取证的原始数据量比较大,还可以进行定位分析,缩小取证的数据范围。

从具体取证程序来看,在云计算环境下进行取证首先需按云端服务类型来确定取证方法。如在云端设备中,客户端信息极有可能在终端短暂存在,此时传统的取证方法仍然可能适用。但是,在云端设备中,现场取证较为困难的数据存取存在一定的不稳定性,且租户存储设备可能会污染采集的数据,不同租户的日志记录可能会设在同一地点或散落在多个不断变化的装置中,故采集数据可能需要大量带宽和时间,数据的碎裂、分散和所有权问题影响数据的采集。在云端软件中,很多信息可以用来取证调查,包括应用程序和验证记录,以及极有可能在终端短暂存在的客户端信息、供应商端设备包含的基本访问信息等。云端软件的功能可能可以协助网络取证。但是,云端软件可能由不同云端服务提供商提供,且其程序可能较为复杂,因此在分析上可能会存在一定的难度。而在云端平台中,可以在客户端取证,以适用于网页服务器或虚拟操作系统的相关取证方法。技术类型是指云计算开发方、客户端和云服务提供商。

其次,云计算环境下进行取证还需要进一步确定其背后技术类型,即云计算开发方、客户端和云服务提供商,并在此基础上针对某技术类型进行具体的取证工作。如对于云计算开发方,收集完整证据,确定上传数据使用的工具,确定与云服务提供商时间的一致性;对于云服务提供商,首先与云服务提供商沟通以保全数据;对于客服端,确定用户在客户端的角色,使用现场取证工具收集在客户端上进行的活动。

3 云计算环境下取证面临的问题

云计算以动态的服务计算为主要技术特征,以灵活的“服务合约”为核心商业特征,是信息技术领域正在发生的重大变革。但这种变革为信息安全领域带来了巨大的冲击,包括在云平台中运行的各类云应用没有固定不变的基础设施,没有固定不变的安全边界,难以实现用户数据安全与隐私保护;云服务所涉及的资源由多个管理者所有,存在利益冲突,无法统一规划部署安全防护措施;云平台中数据与计算高度集中,安

全措施必须满足海量信息处理需求^[13]。

基于云计算存在上述安全风险,针对云计算的犯罪会频繁发生的可能性较高,在云计算环境下展开取证工作,则可能面临如下问题。

3.1 技术方面的问题

云计算环境下进行取证面临的技术问题集中体现在以下两个方面。

3.1.1 难以确保电子证据信息的全面性

数据采集是要在云计算环境下可能的数据源中鉴别、标识、记录和获得电子数据。在云计算环境下,电子数据的存储大部分都是在云中完成,电子数据并非保存在一个确定的物理节点上^[14],即使在用户本地也可能没有完整的数据副本,因此要全面采集完整的犯罪证据就存在一定困难。

另外,在云计算环境下不同客户数据的混合,尤其是保密、敏感数据在储存或传输过程中易与其它客户数据混合,从而导致云计算环境中的数据具有高度复杂性及易被污染性。这将在数据安全和地缘位置等方面增加安全问题,同时也会给所取电子证据的原始性、完整性和有效性带来一定负面影响。

3.1.2 取证手段的欠缺性

在云计算环境下,某些电子证据在目前的技术取证手段下,难以充分侦测与提取。如在云应用服务中,有至少 1/3 的云实例是通过运行的虚拟专用网络(VPN)访问的,这种方式对网络刑侦几乎检测不到;如果 VPN 是通过网吧、图书馆或公众开放环境访问,那么更没有办法侦测到;大量的关键多媒体终端设备通过无线方式接入到开放网络中,都使这个问题更加复杂和恶化;还有其他的操作因素也会造成取证工作的资料缺失,比如人为故意将机房中工作器电源关闭,这实质上可能破坏了在 RAM 中的大量有价值证据,这样任何连接到云的实例都会丢失^[15]。

3.2 法律方面的问题

云计算环境下进行取证工作,其最终目的是要为诉讼服务,因此取证的环节必须依法进行。尽管目前我国有一些规范性文件对计算机取证问题有一定的规制(如《办理死刑案件证据规定》第 29 条),但从总体上来看,相关立法并不完善,这也使得在云计算环境下进行取证尚面临一定的法律问题。

3.2.1 跨管辖权合法取证难度较大

云计算与传统的外包服务不同,其主要区别在于借助云计算,数据通过互联网进行存储和交付,数据的拥有者不能控制,甚至不知道数据的存储位置,数据的流动是全球性的,跨越了国界、穿越了不同的时区。因此在云计算环境下进行数据的处理或存储所使用的分布式系统和它们处理的数据,可能会存在跨区域乃至跨国界的情况。云计算环境下,数据跨境传送、即时性地在全球传播,而不同的国家都拥有自己的法律以及管理要求,因此在取证过程中要充分识别所涉及系统或数据的精确定位,并确定其所属管辖权,从而进行合法取证,确保证据的效力,这一点实现的难度较大。

3.2.2 取证程序的合法性规范不够具体

云计算通过集群应用、网格技术或分布式文件系统等功能,将网络中大量各种不同类型的存储设备通过应用软件集

合起来协同工作,共同对外提供数据存储和业务访问功能。而这个网络是基于本机防火墙之外的存储设备,所以增加了用户对隐私保密性的担忧。因此,在云计算环境下进行取证,不可避免地会面对大量的个人隐私及商业秘密,如分寸把握不当,极有可能造成侵犯他人隐私权或商业秘密的违法后果。而我国现行规范性法律文件则并未考虑到“云”时代取证问题的复杂性,尤其是对取证程序的合法性问题的规定过于原则,该问题的存在必然会给取证人员的取证行为带来一定的挑战。

4 云计算环境下的取证对策

基于云计算环境下取证存在的上述技术及法律问题,可重点从以下几个方面展开探索。

4.1 技术对策

4.1.1 数据迁移技术

数据迁移技术是存储系统的核心技术之一,其在云计算环境下的取证过程中较显重要。具体而言,一方面可以采取面向取证的现场迁移技术,即在云计算平台中通过对虚拟化软件层的控制,对硬件配置和系统服务进行迁移,实现了对操作系统的动态迁移技术的运用。然后通过对虚拟化软件层迁移状态时的上层虚拟机的进程标识、文件系统信息重构、内存信息的迁移、网络连接信息监控,以及注册表信息等进行保存和重构设计,以完整地保存虚拟机制的整个系统状态,并通过本地化镜像加载,将虚拟机镜像整个从云计算平台迁移到本地取证环境中进行分析,实现云计算平台下电子证据的获取。

另一方面还可采取镜像迁移数据的本地化取证,即建立虚拟机重现的模型,借助磁盘虚拟化的动态迁移、动态硬件配置的迁移,以及动态服务和指令的迁移,可以在实际的主机环境中将镜像的虚拟机实例运行起来,实现了本地虚拟化技术,让镜像的虚拟机直接从本地设备的操作系统启动。同时,由于迁移出来的虚拟机镜像文件需要在本地加载才能进一步进行取证分析,因此为了使镜像文件可以正常在本地环境下加载,可以在本地磁盘中重新划分一个磁盘区域作为本地系统和虚拟机镜像系统的信息交互场所,从而促使虚拟机镜像实现本地系统的顺利加载。

4.1.2 入侵检测系统

入侵检测系统可以全面分析检测客户端与云端资源的交互数据,一旦发现有与入侵模式库中相匹配的信息,系统就会做出响应,并作出处理。入侵检测不但可以在检测到危险时发出警告响应,还有一套安全机制保护客户端数据,可以有效防止并减少内部入侵和恶意软件造成的危害^[16]。

4.2 法律对策

首先,健全国家的法律法规和司法制度,建立与云服务提供商的合作,在业务流程上实施监控,获取运行过程的系统信息,留待取证分析^[17]。同时,在程序法的完善方面,则一方面要重视对取证程序的具体规范,明确在云计算环境下进行证据调取时的合法底线,做到既能有效获取有关线索与证据,又要保证取证的合法性,体现正当的要求;另一方面,需对取证技术及取证流程进行细化规定,从而将程序正当性原则具体化,确保取证工作的准确性及有效性。

其次,针对跨管辖权问题的解决,则宜提高取证人员取证

前对当地有关取证法律程序的认识水平,以确保取证活动符合该国或该区域的相关规定;同时还可以积极开展国际(区际)合作,以确保取证活动的顺利进行。

最后,为了适应时代发展的需要,更好地保护网络隐私权,立法机关应尽快加强隐私权方面的相关立法,明确隐私权的法律地位,建立一套完善的网络隐私权保护体系,包括准确界定网络隐私权,建立相应的监管机构等。

结束语 云计算是当前发展十分迅速的新兴产业,也是未来网络的发展趋势。较之传统取证模式,在云计算环境下开展取证活动在资源整合、设备整合、信息共享以及效益最大化等方面具有一定的优势,未来这种取证模式必然会得到进一步的广泛适用。但由于技术、法律、监管等多方面的原因,在云计算环境下所进行的取证活动可能会不断面临新的技术及法律问题。如何有效解决这些问题,充分发挥云计算环境下的取证效益,则成为云计算环境下取证活动研究之主要方向。

参考文献

- [1] 栾润生. 面向云计算的计算机网络犯罪侦查取证的思考[J]. 网络安全技术与应用, 2011, 12: 68
- [2] Boss G, Malladi P, Quan D, et al. Cloud computing. IBM White Paper[EB/OL]. <http://download.boulder.ibm.com/>, 2010-12-10
- [3] 夏荣. 云计算技术在电子数据取证领域的应用研究[J]. 网络安全技术, 2011, 8: 52
- [4] 曹彬, 程久军, 闫春钢. 基于云计算 Live Mesh 的流媒体应用研究[J]. 计算机科学, 2010, 11: 92
- [5] 彭召意, 周玉, 文志强. 基于云计算的视频取证监控系统[J]. 计算机应用研究, 2011, 8: 2975
- [6] 吴同, 杨永川. 云计算环境下的取证研究[J]. 电信科学, 2010, 12: 80
- [7] Goldberg R. Survey of Virtual Machine Research[J]. IEEE Computer, 1974, 7(6): 34-45
- [8] 周刚. 云计算环境中面向取证的现场迁移技术研究[D]. 武汉: 华中科技大学, 2011: 12
- [9] 李小恺. 云计算环境下计算机侦查取证工作研究[D]. 北京: 中国政法大学, 2011: 14
- [10] 丁秋峰, 孙国梓. 云计算环境下取证技术研究[J]. 信息网络安全, 2011, 11: 37-38
- [11] 张超. 云计算环境下的电子数据调查与取证[J]. 信息网络安全, 2010, 10: 54
- [12] 周刚. 云计算环境中面向取证的现场迁移技术研究[D]. 武汉: 华中科技大学, 2011: 29
- [13] 冯登国, 张敏, 张妍, 等. 云计算安全研究[J]. 软件学报, 2011, 1: 74
- [14] 丁秋峰, 孙国梓. 云计算环境下取证技术研究[J]. 信息网络安全, 2011, 11: 37
- [15] 周刚, 麦永浩, 曹强, 等. 云计算运用对计算机取证技术的挑战及对策[J]. 警察技术, 2011, 3: 38
- [16] 丁秋峰, 孙国梓. 云计算环境下取证技术研究[J]. 信息网络安全, 2011, 11: 38
- [17] 宋蕾, 李俊丽. 云计算环境下的计算机取证[J]. 河南科技, 2011, 1(上): 57
- [18] 李德毅. 云计算支撑信息服务社会化、集约化和专业化[J]. 重庆邮电大学学报: 自然科学版, 2010, 22(6): 698-702