# 基于云的计算机取证系统研究

武 鲁 王连海 顾卫东
(山东省计算机网络重点实验室 济南 250014)

**摘 要** 云计算是目前最流行的互联网计算模式,它具有弹性计算、资源虚拟化、按需服务等特点。在云计算的环境中,云计算中心直接提供由基础设施、平台、应用等组成的各种服务,用户不再拥有自己的基础设施、软件和数据,而是共享整个云基础架构。这种方式直接影响了云计算环境的安全性和可用性,给云计算带来了巨大的隐患。在分析和研究云计算环境下的安全缺陷和安全威胁的前提下,提出了一种基于云的计算机取证系统的设计方案——利用计算机取证技术解决云计算环境的安全问题,同时利用云计算技术和超算技术满足传统取证对高性能计算的需求。

**关键词** 云计算,超算,计算机取证
**中图法分类号** TP309 **文献标识码** A

## Research on Computer Forensics System Based on Cloud Computing

WU Lu WANG Lian-hai GU Wei-dong
(Key Laboratory for Computer Network of Shandong Province, Jinan 250014, China)

**Abstract** Cloud computing is the most popular computing mode of Internet, which features elastic computing, resource virtualization and on-demand service, etc. In this environment, resources such as infrastructures, development platforms and applications are directly provided by could computing center. Users no longer own infrastructures, software and data themselves, but share entire cloud computing resources. This will affects security and availability of could computing environment, and leave it tremendous risks. Security flaws and threats of cloud computing were analyzed in this paper, and a design of computer forensic system based on cloud computing was presented, which resolves security problems of cloud computing using computer forensic techniques, and meets the need of high-performance computing using cloud computing and super computing techniques.

**Keywords** Cloud computing, Super computing, Computer forensics

## 1 云计算的安全威胁

云计算是当前IT行业的发展热点,虽然它还没有非常确切的定义,但是云计算的特征是毋庸置疑的。即,云计算中的硬件和软件资源在物理上是分布式的,而在逻辑上是整体呈现的,其资源可以动态地扩展和配置,最终以各种服务的形式提供给用户,用户只需按照实际需求付费使用云中的资源,而不需进行额外的管理。云计算的运行方式导致其在资源层、数据层、应用层上都存在着很多风险和安全威胁,虽然有些并非是云计算特有的,但其特点导致了这些风险和威胁被严重放大。

### 1.1 资源层

在网络资源层面,云计算中的各台物理主机都是通过互联网进行连接的,因此首先要确保互联网资源的可用性和所有网络资源都有适当的访问控制。如果互联网资源失效,整个云就会崩溃。

在主机资源层面,云计算管理了数量巨大的物理主机节点,这些节点采用相同的操作系统、防护措施和安全策略,这就意味着风险被放大了成千上百倍:如果一台物理主机遭到破坏,整个云计算的物理主机都可能被破坏,物理主机中的所有虚拟机以及虚拟机提供的服务也将不复存在,这被称为云计算的"高速攻击"。虚拟化技术的安全漏洞也被带入云计算当中,例如用户可能利用虚拟机的漏洞,将恶意代码注入到物理主机系统中,以此来控制物理主机或其他虚拟主机,这被称为"虚拟机逃逸"[1];虚拟化管理程序的漏洞可能会导致虚拟机之间的隔离不彻底,例如有的管理程序使用共享剪贴板以方便虚拟机和物理主机之间的通信,但是其他虚拟机可以通过堆栈溢出的方式访问到共享剪贴板[2]。另外,在虚拟化环境下CPU的负载、CPU Cache以及磁盘和内存的一些共享机制也会产生隐蔽信道[9],造成虚拟机信息泄漏。

### 1.2 数据层

数据的安全威胁包括两种:一是数据的安全性问题,即用户数据不被攻击和篡改;二是保密性问题,即用户的数据不被未授权的人访问。虽然云计算数据的大规模分布式存储机制

到稿日期:2011-06-17 返修日期:2011-09-19 本文受国家自然科学基金(61070163),山东省自主创新成果转化重大专项(2011ZHZX1A0109)资助。

武 鲁(1982-),男,硕士,助理研究员,主要研究领域为计算机取证、云计算,E-mail:Wul@keylab.net;王连海(1969-),男,硕士,研究员,主要研究领域为计算机取证、网络信息安全、云计算;顾卫东(1970-),男,硕士,研究员,主要研究领域为云计算、超算、信息安全。

增加了非法访问数据的难度，但是黑客仍可以通过暴力破解所有存储服务器来收集分块信息，甚至可以破解云存储管理系统的数据分发逻辑，从而精准地定位每一个数据块。另外，为了便于索引和查询，分布式存储系统中的数据都是不经过任何加密的，所有用户的数据都混在一起存放，应用程序设计中虽然包含了防止对未授权数据的访问数据标签，但是仍有可能利用应用程序的漏洞实现未授权访问。

### 1.3 应用层

应用层的安全威胁主要集中在利用网络应用程序（如ERP、CRM、OA、门户网站、博客、论坛等）的漏洞和错误对其进行攻击，如 SQL 注入、执行恶意代码等。另外，应用层的 DoS 和 DDoS 攻击（如网页重新加载、XML Web 服务请求、云计算服务支持的特定协议的请求等）也会快速消耗云计算的资源。目前，在 IaaS 平台的服务过程中，仅能实现对资源的简单监控，而对应用层没有适当的审计和监控手段，因此无法保证其不被误用和滥用，不久的将来可能会出现从云计算中发起的对其他云计算服务的攻击。

## 2 计算机取证在云计算中遇到的难题

计算机取证（Computer Forensics）是利用各种计算机软硬件知识和辨析技术，对计算机入侵、破坏、攻击、欺诈等犯罪行为，按照符合法律规范的方式进行识别、保存、分析和提交数字证据的过程[7,8]。计算机取证一般分为离线和在线两种方式（又称为静态取证和动态取证）。离线取证是指关闭计算机或电子设备后进行的证据获取、分析和呈现过程，主要针对磁介质存储器和光存储器中的非易失性数据，如文本、数据库、音视频等多种格式的文件。在线取证是指在不关闭计算机或电子设备的情况下进行的证据获取、分析和呈现过程，尤其是通过物理内存分析的在线取证方式来获取系统进程信息、加载的驱动程序、网络连接状况、当前打开的文件、其他系统操作痕迹等大量易失性数据[3-6]。

计算机取证技术是打击计算机网络违法犯罪活动、保障信息安全的最直接的武器，它是网络安全技术的一个重要分支，不仅可以用于网络安全防护和应急响应，必要时还可以介入司法程序，为法律诉讼提供可采信的证据。然而在云计算环境下，传统计算机取证遇到了很多新挑战。

首先，传统的离线取证在大规模分布式存储环境中已经失效。传统的离线取证方式必须直接对目标计算机的磁盘进行克隆来收集犯罪证据，但是在分布式存储环境中却不可行。第一，一个完整的文件会被分割成若干数据块，并存储在不同的节点上，而各节点可能存在于不同的地域，存储的非定域性和司法管辖权的限制导致取证的复杂程度和成本激增；第二，离线取证要求云计算的分布式存储系统节点全部或部分停机，这对云计算服务来说是不可接受的。

其次，单机无法解决海量数据的存储、分析问题。目前计算机取证分析的主要依靠是单台计算机（或工作站），取证分析的过程就是从原始证据中剔除无用数据，找到有用的证据。但随着存储器容量的迅速增大，对云计算进行取证将产生比对单机取证高出几个数量级的海量数据，这些数据的存储、分析、检索需要海量存储空间和强大的计算能力，这都是单台计算机无法完成的。

第三，取证对高性能计算的需求。密码和口令问题一直是计算机取证中经常遇到难题之一，在云计算时代，取证对口令密码破解的时效性要求更高。在加密算法和安全认证体系不存在巨大漏洞的情况下，穷举已成为口令密码破解的唯一方式，但是穷举所产生的计算量也是普通计算技术无法解决的。

## 3 基于云的计算机取证系统的设计

针对上述云计算的安全威胁和计算机取证遇到的难题，基于云的计算机取证系统在设计思路上基于以下两点：一是利用各种计算机取证技术来解决云计算环境下的安全问题；二是将云计算和超算的高性能计算能力和大规模分布式存储环境应用于计算机取证分析，来解决云计算取证中遇到的各种问题。本系统可以对云计算环境中的各节点（物理主机、虚拟机、客户终端、移动终端等）进行取证，它包括云取证服务中心、客户端取证工具和云取证中间件 3 大部分。

### 3.1 云取证服务系统

云取证服务中心包括取证分析单元、证据存储单元、超算处理单元、管理单元 4 个部分内容，如图 1 所示。

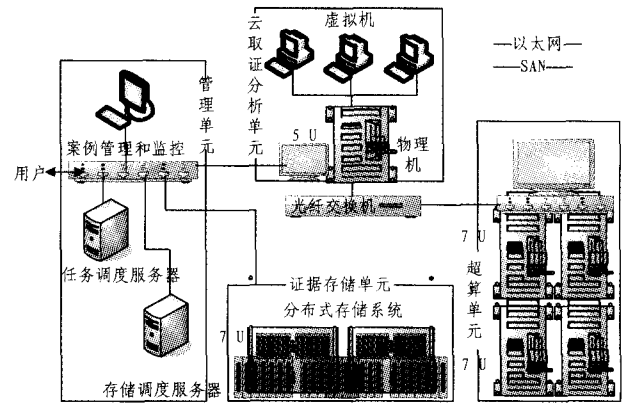


图 1 云取证服务中心体系结构

1）管理单元：实现云端的管理功能，包括：

①任务调度：完成取证分析和证据存储的任务调度工作。根据计算量和数据量，利用云计算中心提供的 IaaS 服务来配置、创建、收回虚拟机，在证据处理过程中动态地完成计算任务分配和虚拟机迁移等调度工作。

②取证控制：接受并审核用户提交的取证请求，远程调用适当的取证工具对目标计算机进行取证。

③案例管理：对进入云端取证服务器的取证分析任务进行的全过程进行监视、审计和记录，以便需要之时向法庭提交，以说明取证流程的合法性。

④数据传输：负责各单元和各模块之间的控制消息和数据传递。

2）取证分析单元：利用管理单元提供的虚拟机，并行处理多个取证分析任务。为超算中心的并行计算环境提供接口，当出现高性能计算的取证分析任务时，直接交予千万亿次超级计算机处理。

3)证据存储单元:用于存储客户端取证工具上传的原始证据和云取证中间件采集的用户行为审计数据,证据分析单元的存储空间由云计算中心的云存储提供。

4)超算处理单元:超算处理单元由国家超算济南中心完成,该中心采用千万亿次计算机"神威蓝光"作为业务主计算机,其上面运行了密码口令分析程序和其他经过并行化移植的计算机取证分析处理程序,用以完成取证分析所需要的高性能计算。

### 3.2 客户端取证工具

客户端取证工具由多款工具组成,用户通过本地下载或远程控制下载的方式将需要的客户端下载到目标计算设备上(物理主机、虚拟机、客户终端、移动终端等),通过本地操作或云取证服务中心的远程控制来完成存储器(磁盘和内存)证据获取工作。证据获取完成后,简单任务可以在本地独立处理,复杂任务可通过云取证中间件将证据推送至云取证服务中心处理。客户端取证工具包括针对计算机和智能终端两大系列。

### 3.3 云取证中间件

云取证中间件是嵌入在客户端操作系统中的进程或系统服务,它屏蔽了用户操作系统和硬件结构的差异,使云取证服务中心可以和运行在客户操作系统上的客户端取证工具以及其他第三方应用程序进行交互。云取证中间件具有以下功能:

1)传递远程调用。云取证服务中心可以根据用户的请求远程控制取证工具和其他第三方应用,以实现证据的实时收集;客户端取证工具和其他第三方应用也可以调用云计算服务中心上的过程实现本地证据的远程处理(如内存分析、证据保存等)。

2)进行安全审计。云取证中间件可以实时获取客户端操作系统的安全审计信息(行为审计、网络流量审计和内容审计),并将其发送到云取证服务中心进行分析处理,以及时发现网络安全隐患或违法犯罪活动。

## 4 基于云的计算机取证系统的运行流程

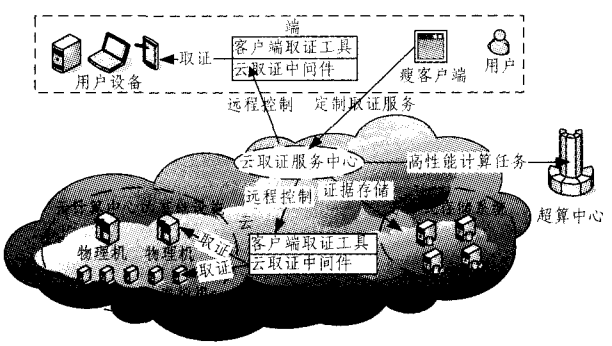


图 2 基于云的计算机取证系统运行方式

第一步,云取证服务中心面向公众提供 SaaS 模式的"云取证服务",用户通过互联网瘦客户端(如浏览器)向云取证服务中心定制取证服务,发出针对某一对象的取证请求。

第二步,云取证服务中心接受并审核用户的请求,控制客户端取证工具和云取证中间件在计算设备上收集原始证据,并将原始证据存放到分布式存储系统上。

第三步,云取证服务中心的任务调度模块利用云计算中心的 IaaS 服务创建适当数量的虚拟进行取证分析处理,对于高性能计算任务(如口令密码分析等)交由超算中心处理。

第四步,将取证分析的结果反馈给用户。

整个运行流程过如图 2 所示。

**结束语** 本系统提出了"云取证服务"的概念,将计算机取证作为一种服务通过云计算的方式提供给用户,再根据用户的请求对各种计算资源(物理主机、虚拟机、移动终端)中的安全威胁、违法犯罪行为进行灵敏感知、事中取证和全程监控,改变了以往事后取证的局面,利用云计算技术和超算技术解决了传统取证中的一些难题,对提高云计算环境的安全性有很大帮助。

云计算的取证还刚刚起步,今后我们将在云计算取证的理论、技术和应用上继续深入研究。首先,要对比传统的计算机取证模型和证据可信性评估架构,给出一个针对云计算的、可信的计算机取证模型。其次,要研究传统的计算机取证技术与云计算的结合方式,并将现有的计算机取证工具进行虚拟化、并行化移植,使其可以满足云计算和超算的运行环境。最后,针对今后云计算中即将出现新式的安全威胁和违法犯罪行为,寻求新的对策。

## 参 考 文 献


[1] Reuben J S. A survey on virtual machine security[R]. Helsinki University of Technology, October 2007

[2] Kirch J. Virtual machine security guidelines[R]. The center for Internet Security, September 2007

[3] Wang Lian-hai, Xu Li-juan, Zhang Shu-hui. A Method on Extracting Network Connection Information from 64-bit Windows 7 Memory Images[J]. 中国通信, 2010, 17(6): 44-51

[4] Zhang Shu-hui, Wang Lian-hai, Zhang Lei. Extracting windows registry information from physical memory[C]// International Conference on Computer Research and Development (ICCRD). 2011, 2: 85-89

[5] Wang Lian-hai, Zhang Rui-chao, Zhang Shu-hui. A Model of Computer Live Forensics Based on Physical Memory Analysis [C]// ICISE'09. Dec. 2009

[6] Zhang Rui-chao, Wang Lian-hai, Zhang Shu-hui. Windows Memory Analysis Based on KPCR[C]// 2009 Fifth International Conference on Information Assurance and Security. 2009: 677-680

[7] 许榕生,吴海燕,刘宝旭. 计算机取证概述[J]. 计算机工程与应用, 2001, 114(21): 7-8

[8] 钱桂琼,杨泽明,许榕生. 计算机取证的研究与设计[J]. 计算机工程, 2002, 28(6): 56-58

[9] 丁丽萍. 云计算环境下隐蔽信道分析关键技术研究[R]. 北京:中国科学院软件研究所, 2011