

云计算数据安全

聂雄丁¹, 韩德志^{1,2}, 毕坤¹

(1. 上海海事大学信息工程学院, 上海 201306; 2. 华南理工大学计算机科学与工程学院, 广州 510641)

摘要: 提出了一种可以同时解决云环境下数据的隐私性、可用性和完整性安全挑战的方法。综述了当前具有代表性的数据安全技术,着重分析了它们的功能单一性缺陷,并通过改进使之更加完善。提出了使用第三方审计不仅可以批量验证云租户数据的完整性,还可以构建云环境的可信链以及保障云租户数据隐私性的新思路。

关键词: 云计算; 数据安全; 第三方审计; 隐私; 分级保护

中图分类号: TP309 **文献标志码:** A **文章编号:** 1671-5497(2012)Sup. 1-0332-05

Cloud computing data security

NIE Xiong-ding¹, HAN De-zhi^{1,2}, BI Kun¹

(1. College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China; 2. School of Computer Science and Engineering, South China University of Technology, Guangzhou 510641, China)

Abstract: A comprehensive method was explored to handle data privacy, availability and integrity challenges in cloud computing. The existing data security techniques were stated, and their solo-function drawback was analyzed and improved. As a consequence, a new idea was proposed to tackle both data privacy and integrity issues by the third party auditor. The proposal not only can verify users' data integrity by batch, but also can construct a credible chain between each party and protect the tenants' data confidentiality.

Key words: cloud computing; data security; third party auditor; privacy; hierarchical protection

当前,云计算发展面临很多关键问题,特别是数据安全问题尤为严重。最近趋势科技所做的一项全球云计算安全调查指出,将近一半(43%)的企业IT决策者表示,自己使用的云计算服务供应商在2011年2月至2012年2月的12个月中曾经出现安全疏漏或安全问题。而近来,Amazon、Google等云计算发起者不断爆出的各种安全事故,更加剧了人们的担忧。因此,要让企

业和组织大规模地使用云计算技术与平台,放心地将自己的数据交付云服务商管理,就必须着手解决云计算面临的各种数据安全问题。本文在分析当前云计算面临的各类数据安全挑战及其解决方案的基础上,提出了使用第三方审计(TPA)来保证云租户数据隐私性的新思路,以期为我国未来的云安全科研、产业发展作出有益的探索。

收稿日期: 2012-05-24.

基金项目: 国家自然科学基金项目(61070154); 上海教委科技创新项目(20110546); 广州市2009难题招贤项目; 广东省科技攻关项目(2010B090400160); 中国博士后科学基金项目(20110490091); 上海海事大学科研基金项目(20110014).

作者简介: 聂雄丁(1987-),男,硕士研究生。研究方向:云安全。E-mail:357481280@qq.com

通信作者: 韩德志(1966-),男,教授。研究方向:云存储,云安全。E-mail:dezhihan88@sina.com

1 云计算数据安全挑战

虽然云服务提供商可以组织安全服务队伍对整个系统进行专业化安全管理,然而由于云计算系统的巨大规模以及前所未有的开放性与复杂性,其安全性面临着比以往更为严峻的考验。表 1 显示了中小型企业用户最担忧的几种安全问题,是根据 2010 年欧洲网络和信息安全局(ENISA)提供的一份调查报告得出的^[1]。

表 1 主要的数据安全问题

Table 1 Main concerns of data security

安全问题	重要/%	很重要/%	合计/%
企业数据的隐私性	30.9	63.6	94.5
服务/数据的可用性	47.3	40.0	87.3
服务/数据的完整性	42.6	44.4	87.0
服务/数据失去控制	47.2	28.3	75.5
服务商缺乏责任感	43.1	29.4	72.5

1.1 数据的隐私性挑战

数据的隐私性是指数据不能被非授权者、实体或进程利用或泄漏的性质。目前,云计算中数据的隐私性是一个不可忽视的问题,它贯穿于数据的存储、传输和消亡过程中。影响数据的隐私性因素主要有:隔离风险、管理接口风险、数据保护、不安全或不完全的数据删除以及内部恶意攻击等。就隔离风险而言,虚拟机的安全性问题就值得注意。虚拟机技术是云计算的重要组成部分,在一台物理存储设备上建立多个虚拟机,把虚拟机分配给多个用户使用,也就是说多个用户的虚拟机可能运行在同一台物理存储设备上,如果虚拟机软件本身存在安全漏洞,那么用户的数据可能被运行在同一台物理存储设备上的其他用户访问。2009 年 5 月,网络上曝光的 VMWare 虚拟化软件的 Mac 版本中存在一个严重的安全漏洞,别有用心的人可以利用该漏洞执行恶意代码^[2]。

1.2 数据的可用性挑战

数据的可用性指数据不会因为恶意攻击等行为而变得不可用,这是用户关心的一个重要问题。影响数据的可用性因数主要有:云服务商基础架构的可靠性和按需计算与 DDoS 攻击等。就云服务商基础架构的可靠性而言,一方面用户想确定云服务商的系统是否 100% 安全可靠,能否为用户提供全天候、不间断的服务,使得用户能够随时随地访问自己的数据;另一方面,某些云厂商提供

的服务是建立在其他云服务商提供的基础架构上,其可靠性还依赖于其他多家云厂商的可靠性,使得其可靠性和可用性问题变得更加复杂,用户数据的可用性与可靠性也受到极大挑战。

1.3 数据的完整性挑战

数据的完整性是指数据没有遭受非法篡改或删除,确保其真实性和有效性。维持云环境下数据的完整性意义重大,它不仅代表了云租户的信息资产安全和利益,而且直接影响云服务商的声誉。影响数据完整性的因素主要有:①从目前的硬盘驱动(或者固态盘、磁带)的发展趋势看,它们的容量增长速度落后于数据的增长率。因此,为了满足云计算海量数据存储的需求,云服务提供商需要不断增加硬盘驱动的数量,这样就极有可能造成节点失效、磁盘失效甚至是数据崩溃或丢失。②尽管磁盘驱动(或固态盘)的容量越来越大,但是并没有提供更快的数据访问速度,有可能造成数据更新失败或数据存取错误等。③黑客等未授权组织或个人对用户数据的篡改或删除等。

2 已有的解决云计算数据安全方案

2.1 关于数据的隐私性

2.1.1 相关研究及实践

Amazon S3 主要在数据的传输和存储阶段都采用加密机制,并提供 4 种访问控制机制:身份和访问管理策略(IAM)、访问控制列表(ACLs)、桶策略、查询字符串认证^[3]。文献[4]建议用户应将自己的数据按重要性等级划分为公有数据和私有数据,不太重要的数据(即公有数据)存放在公有云中,而对敏感数据(即私有数据)则需要构建私有云或者混合云来实现弹性计算和数据隐私的均衡,同时也为未来公有云平台上的实施积累经验。

2.1.2 虚拟机的安全问题

可以使用虚拟机扫描技术解决虚拟机的安全问题。通过扫描虚拟机本身和虚拟机中安装的软件,确保当前用户的虚拟机运行正常,没有进行非法的计算和访问。惠普公司和 IBM 公司均提出了自己的虚拟机扫描技术,其原理见图 1^[5]。

云服务提供商在每台物理上除了为客户建立和提供虚拟机(VM)外,还建立了一个用于安全监控的虚拟机供云服务商自己使用。该安全监控虚拟机可以随时扫描其他的虚拟机,分析其运行状态和行为,而其他虚拟机无法感觉到该虚拟机



图 1 虚拟机扫描技术图

Fig. 1 VM-scan technology map

的存在,因此也无法阻止或发现安全监控虚拟机的扫描行为。通过这种技术,很多公共的扫描需求都可以在该安全监控虚拟机中完成,而不需要在用户自己的虚拟机中安装扫描软件,一方面简化了扫描操作,更重要的是,用户无法干扰虚拟机安全扫描软件的工作。

2.2 关于数据的可用性

解决数据的可用性问题的一般方法是冗余备份。Gartner 建议,在选择云服务商时企业用户不但需要了解云服务商是否具备数据恢复的能力,而且还必须知道云服务商能在多长时间内完成数据恢复。Hadoop 的分布式文件系统 HDFS 采用机架感知(Rack-aware)的策略来改进数据的可靠性、可用性和网络带宽的利用率^[6]。通过机架感知,NameNode(管理文件系统的元数据)可以确定每个 DataNode(存储实际的数据)所属的机架 ID。一般情况下,当复制因子是 3 时,HDFS 的部署策略是将一个副本放在同一个机架上的另一个节点,一个副本存放在本地机架上的节点,最后一个副本放在不同机架上的另一个节点。机架的错误远比节点的错误少,这个策略可以防止整个机架失效时数据丢失,不会影响到数据的可靠性和可用性,又能保证性能。

2.3 关于数据的完整性

2.3.1 一般解决方案

保证数据完整性通常有两种方法:使用信息认证编码(MAC)和数字签名(DS)^[7]。MAC 中,依赖对称密钥产生校验和附加在数据后面,而 DS 算法则依赖公共密钥结构(有公钥和私钥对)。由于对称算法比非对称算法快,所以认为 MAC 是最好的完整性检查机制。目前,一般使用数字签名来进行数据完整性测试,例如广泛采用的分布式文件系统(如 GFS、HDFS)将大的数据卷划分为若干默认大小(64 MB 或 128 MB)的数据块,每个数据块都会附上一个数字签名物理存储起来,用于日后的完整性测试^[8]。

2.3.2 引入第三方审计(TPA)的解决方案

云租户亲自进行数据完整性验证会面临计算资源不足、任务繁琐、密钥管理等问题,另外,即便

检测出完整性问题也难以确定“肇事者”。在这种情形下,为用户和云服务商双方所信赖的第三方审计(TPA)应运而生。TPA 必须满足支持数据动态更新、公共审计和安全审计等要求^[9]。

TPA 数据完整性审计过程如下:①用户向 TPA 发出数据审计请求;②TPA 向云服务商提交用户审计请求;③云服务商向 TPA 返回相关数据信息;④TPA 审计云服务商返回的数据信息;⑤TPA 向用户返回数据完整性审计结果。

3 分析与改进

3.1 存在的问题

随着云计算的不断普及,针对云安全问题的相关研究也越来越深入。但基于云计算前所未有的开放性和复杂性,依靠目前的安全技术,并不能解决所有安全问题。例如,在 TPA 模式中,当用户数据被非法篡改、删除、丢失时,TPA 可以及时准确地向用户报告数据的完整性状态,避免用户损失的进一步扩大,为用户决策提供参考信息。但是,如果云服务商或者其他未授权个人或组织并未破坏用户数据的完整性,而只是非法访问或读取用户数据,TPA 并不能向用户反馈这一信息。现阶段云环境中数据的隐私性完全由云服务商保证,而云服务商部署的安全措施则主要针对黑客等未授权人员,如何有效应对云服务商非法访问或读取用户隐私数据还缺乏相关研究。

3.2 解决思路

在现有的 TPA 模式下引入证书管理可以有效保证数据的机密性。由 2.3.2 节可知,TPA 一方面保证用户可以随时检查数据的完整性,另一方面也能确保数据隐私信息对 TPA 的透明性。如果将用户证书、SaaS 证书、PaaS 证书、IaaS 证书以及硬件层证书都托付 TPA 管理,并保证 TPA 对证书的透明,那么就可以构建用户端到云计算硬件层的安全可信链^[10]。这不仅维持了 TPA 原有的数据完整性检测功能,而且可以有效遏制云服务商或未授权人员非法访问或读取用户数据的行为。

为了确保用户数据的机密性,规定:云环境下的数据访问请求,必须是由合法用户发出的,若访问者的 TPA 身份认证不能通过,则本次访问失败,TPA 向用户报告数据非法访问事件。TPA 身份认证过程如下:①用户首先将自己的个人证书送给 TPA 认证;②认证通过后,相应的云应用

服务商也需要将自己的证书提交认证;③同样地, IaaS 和硬件层只有等上层认证通过后, 才能向其下层发出访问请求。图 2 是用户身份认证过程和数据访问过程。

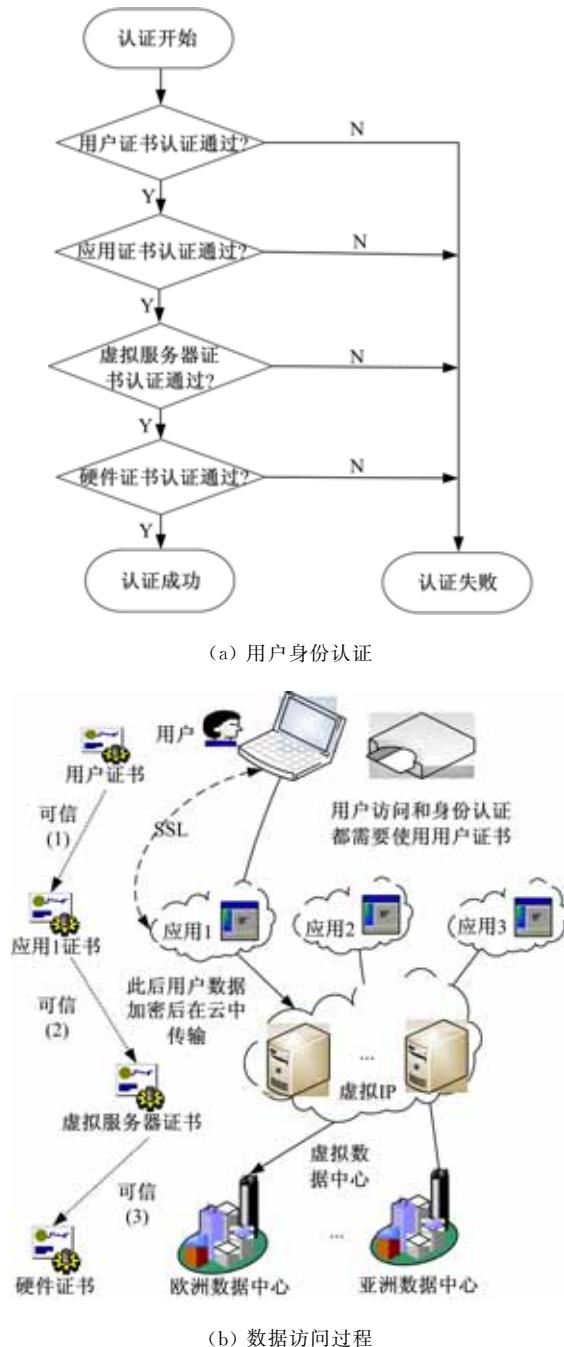


图 2 用户身份认证和数据访问过程

Fig. 2 User identity authentication and data access procedure

特别需要说明的是, 若云服务商想要读取某个用户的 data, 而该用户却没有发出数据访问请求, 一方面, 云服务商 TPA 身份认证不能通过, 读取用户 data 失败; 另一方面, TPA 向用户报告

云服务商非法读取自己 data 的行为。通过这种方法, TPA 同时具备了 data 完整性和隐私性检测功能, 大大缓解了用户担心 data 失去控制的焦虑。

为了尽可能降低用户 TPA 开销, 建议用户将云中的 data 按重要程度划分等级, 即用户自定义 data 的安全等级。对安全等级高的 data, 可以申请 TPA 全天候报告 data 隐私性和完整性状态, 而对一般 data 则采取按月或按年报告的方式。在云服务商方面, 用户也可以采取这种方式申请对 data 进行有重点、多层次地保护。文献[11]提出了根据 data 的隐私性(C)、完整性(I)和可用性(A)划分保护环的思想, 这样 data 的 CIA 值反映了 data 的重要性程度。具体步骤为: ① 分别计算某个 data i 的机密性值 $C[i]$ 、完整性值 $I[i]$ 以及可用性值 $A[i]$; ② 计算 data i 的 CIA 值 $CIA[i] = (C[i] + (1/A[i]) * 10)/2$; ③ 判断 $CIA[i]$ 所属的保护环 $R[i]$, 如果 $CIA[i] = \{1, 2, 3\}$, 则 $R[i] = 3$, 如果 $CIA[i] = \{4, 5, 6\}$, 则 $R[i] = 2$, 如果 $CIA[i] = \{7, 8, 9\}$, 则 $R[i] = 1$; ④ 如果 $R[i] = \{x | x = 1, 2, 3\}$, 则将 data i 存入保护环 x 中。保护环结构图如图 3 所示。

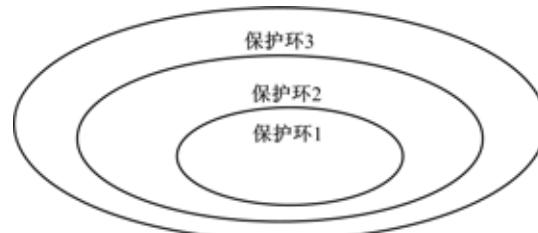


图 3 CIA 保护环结构

Fig. 3 Architecture of CIA protection ring

4 结束语

云计算是当前发展十分迅速的新兴产业, 具有广阔的发展前景, 但同时其它所面临的安全技术挑战也是前所未有的, 需要 IT 领域与信息安全领域的研究者共同探索解决方案。同时, 云计算 data 安全不仅仅是技术问题, 它还涉及标准化、监管模式、法律法规等诸多方面。因此, 仅从技术角度出发探索云计算 data 安全问题是不够的, 需要学术界、产业界、政府等相关部门共同努力才能实现。

参考文献:

- [1] Hogben G. Privacy, security and identity in the cloud[C] // 8ENISA. OASIS/EEMA eIdentity Con-

- ference, London, 2010.
- [2] 冯登国, 张敏, 张妍, 等. 云计算安全研究[J]. 软件学报, 2011, 22(1): 71-83.
Feng Deng-guo, Zhang Min, Zhang Yan, et al. Study on cloud computing security[J]. Journal of Software, 2011, 22(1): 71-83.
- [3] Amazon simple storage service (AmazonS3) [EB/OL]. [2012-03-28] <http://aws.amazon.com/de/s3/>. 2012-03-28.
- [4] Yang Jian-feng, Chen Zhi-bin. Cloud computing research and security issues[C]// International Conference on Computational Intelligence and Software Engineering(CiSE), Wuhan, 2010.
- [5] Anthes Gary. Security in the cloud[J]. Communications of the ACM, 2010, 53(11):16-18.
- [6] 刘鹏. 云计算[M]. 2 版. 北京: 电子工业出版社, 2011.
- [7] Almulla Sameera Abdulrahman, Yeun Chan Yeob. Cloud computing security management[C]// Second International Conference on Engineering Systems Management and Its Applications(ICESMA), Sharjah, 2010.
- [8] Zhou Min-qi, Zhang Rong, Xie Wei, et al. Security and privacy in cloud computing: a survey[C]// Sixth International Conference on Semantics Knowledge and Grid(SKG), Beijing, 2010.
- [9] Hao Zhuo, Zhong Sheng, Yu Neng-hai. A privacy-preserving remote data integrity checking protocol with data dynamics and public verifiability[J]. IEEE Transactions on Knowledge and Data Engineering, 2011, 23(9): 1432-1437.
- [10] Zisis Dimitrios, Lekkas Dimitrios. Addressing cloud computing security issues[J]. Future Generation Computer Systems, 2012, 28(3): 583-592.
- [11] Prasad Parikshit, Ojha B, Shahi R R, et al. 3 dimensional security in cloud computing[C]// 3rd International Conference on Computer Research and Development (ICCRD), Shanghai, 2011.