



基于分离密钥的云存储加密解决方案^{*}

王志中¹, 周 城², 牟宇飞¹

(1. 解放军重庆通信学院 重庆 400035; 2. 重庆市信息安全重点实验室 重庆 400035)

摘 要: 原有的云存储模式所存在的问题已经困扰行业多年, 在静态数据加密存储的过程中, 只有实现真正意义上的数据私有化才能保证数据的安全, 保障数据拥有者的利益。针对云存储应用中用户数据安全存储需求的提高, 以 Amazon S3 (simple storage service) 为例分析当前云存储模式下静态数据安全存在的普遍问题, 设计了一套新的云端静态数据加密存储方案——分离密钥存储服务 (separated key S3) 解决方案, 设计了新的密钥管理方法, 从技术上实现了云端静态数据的安全, 最后对该方案的数据安全性进行了分析。

关键词: 云存储; 分离密钥; 静态数据; 数据安全; 主密钥

doi: 10.3969/j.issn.1000-0801.2013.01.008

A Project Designed by Cloud Storage Encryption Based on Separated Key

Wang Zhizhong¹, Zhou Cheng², Mou Yufei¹

(1. Chongqing Communication Institute, Chongqing 400035, China;

2. Chongqing Key Laboratory of Information Security, Chongqing 400035, China)

Abstract: The problems existing in the original cloud storage model has plagued the industry for many years, to realize the true sense of the data privatization in the static data encrypted storage process is able to ensure data security and protect data owners' interests. Cloud storage applications for the current secure storage of user data is able to improve data security. Taking Amazon S3 as an example, the general problem in static data security was analyzed. A new set of cloud static data encryption and storage solution, separated key S3, was designed, and a new key management method was also designed. The new technology absolutely accomplished the security of static data in the cloud. Additionally the data security analysis and prospect on next research was given.

Key words: cloud storage, separated key, static data, data security, master key

1 引言

云存储是未来信息存储的一种理想方式, 它提供方便易用的外包存储空间, 但是, 由于云存储客户担心自己的数据丢失, 缺乏对云存储的信任, 导致云存储推广普及变得困难^[1]。客户是否放心将私密数据存放到云中, 是云存储是否能够得到广泛推广的关键所在, 而客户的信任感并不

能简单依靠道德上的约束或管理上的规则, 如果能够从技术层面上实现客户数据的完全私有化, 就能从根本上解决客户对云存储的信任问题。从目前诸多实例分析, 云端数据的泄露, 从表面上看, 是源于一些得到操作资源特权的人, 如服务器管理人员或成功入侵服务器的黑客; 从根源上来看, 现在广泛存在的行业标准化, 如编码机制、通用方法 (如通用的加密解密方法) 等才是产生数据泄露的根源。

^{*} 国家自然科学基金资助项目 (No.61272043)



因此,本文设计了一种面向云存储的静态数据存储系统模型及密钥管理策略,旨在完成客户数据的完全私有化,实现云存储的安全。

2 云存储应用中的加密技术

云存储应用中的存储安全包括认证服务、数据加密存储、安全管理、安全日志和审计^[2]。对用户来说,在上述4类存储安全服务中,存储加密服务尤为重要。加密存储是保证客户私有数据在共享存储平台的核心技术,是对指定的目录和文件进行加密后存储,实现敏感数据存储和传送过程中的机密性保护。根据形态和应用特点的差异,云存储系统中的数据被分为两类:动态数据和静态数据。动态数据是指在网络中传输的数据,而静态数据主要是指存储在磁盘、磁带等存储介质中的数据^[3]。

2.1 云数据加密模式

2.1.1 静态数据加密

对于那些在云中存储备份自己长期数据的客户,他们可以将自己的数据加密,然后发送密文到云存储提供商(cloud storage provider, CSP)。目前,大多数该类解决方案基于用户的数字证书进行认证和加密。用户使用数字证书向云管理系统进行身份认证,并使用对称密钥在本地加密云中存储的数据,同时使用证书公钥加密对称密钥,然后将加密后的数据传到云中进行存储。这些客户控制并保存密钥,当客户要获取数据时,先将云中密文下载到本地,再由客户自行解密该数据。

该模式的优点是:只有客户可以解密云中的存储数据,可以有效地保证数据的私密性。缺点是:客户端需要较强的密码运算能力来实现加密功能,同时用户数据的加密密钥必须保管安全,一旦丢失,将无法恢复数据,且该模式只适用于客户自己生成的静态数据加密,对于在IaaS、PaaS、SaaS下在云中产生的动态数据无法使用该模式进行加密^[4]。

2.1.2 动态数据加密

对于在云计算环境中产生的动态数据,只能由云计算管理系统进行加密,虽然CSP的网络比开放网络安全。在多租户的云计算应用模式下,客户租用云计算系统的计算能力,虚拟化技术使得一个客户的应用以不同层次与其他客户的应用共享物理资源,因此客户在云计算环境中产生的数据不可避免地交由云计算环境进行加密。

该模式的优点是:客户不需要为不同的云应用保管各种不同的密钥,而是交由云计算环境统一管理,具有更高

的安全性。缺点是:客户无法控制动态数据,一切依赖于云计算服务提供商,同时云计算管理系统需要提供一个统一、有效、可扩展的云计算密钥管理框架,用于为各类客户提供各种类型密钥的统一管理,实现各种密钥操作。

2.2 数据加密存储技术研究现状

目前云存储加密的研究较多,参考文献[1]提出一种面向云存储的安全存储模型及存取策略,参考文献[5]中给出一种基于层次属性加密实现的细粒度访问控制的云存储系统,参考文献[6]描述了基于加密策略增强云存储的安全性,参考文献[7]提出了一种云存储密文策略属性基加密方案,但此类文献所采用的方式均为上文所述的客户加密数据模式,无法克服该方法本身所具有的缺点。据目前可检索到的文献分析,还没有一种静态数据加密方法既突破客户加密数据上传的传统云存储模式,又能较好地保证数据私有化实现静态数据的安全。

静态数据一直是通过公钥基础设施(PKI)技术保护。在IaaS环境中,使用多种提供商和第三方工具加密静态数据很普遍。在PaaS环境中加密静态数据一般会较复杂,需要提供商提供的或专门定制的设备^[8]。在SaaS环境中加密静态数据是客户无法直接实施的,需要向他们的提供商请求。2011年8月来自微软的安全研究员Kristin Lauter和同事为基于同态加密的加密技术研究出存储原型,该技术的产生及在云端领域的应用,将极大提升云端数据安全性,除了存放者本人外,他人无法破译^[9]。初创安全公司Porticor在2012年3月推出了一个解决方案,专门解决云环境中静态数据安全问题,该公司提供了一种加密解决方案,并声称云客户是唯一知道主密钥的人。

3 Amazon S3 模式云存储加密解决方案

保证云端静态数据安全的最大挑战在于密钥管理。网络安全传输过程中加密数据的存在往往只有几毫秒到几秒的时间,但是存储系统中的加密数据需要保存的时间可以到几年甚至上百年,如果加密系统的实现方式不当或加密算法的强度不够,会导致攻击者有充裕的时间用于尝试不同的攻击手段,极大增加存储系统的风险。下面就以主流云存储服务Amazon S3 (simple storage service, 简单存储服务)^[10]模式为例进行分析。

3.1 Amazon S3 云存储服务的数据加密及密钥管理方案

Amazon公司是最早推出云存储服务的企业,也是最成功的企业。为了利用闲置的硬件资源,Amazon公司从

2006 年开始对外出租云存储服务,即 S3,该服务是 Amazon 网络服务 (Amazon web service, AWS) 的一部分。Amazon 网络服务由 4 个部分组成:简单排列服务 (simple queuing service)、简单存储服务 (simple storage service)、弹性计算云 (elastic compute cloud, EC2) 和简单数据库 (simple DB)。由于存储费用低廉、服务稳定, S3 获得了大量客户的青睐。此外, Amazon 还推出了弹性块存储 (elastic block storage, EBS) 技术, 支持数据的持续性存储。其中 S3 和 EC2 标志着云计算发展进入新阶段^[1]。

图 1 是 S3 的工作原理示意。如图 1 所示, S3 服务数据存储方案如下。

(1) 客户首次登录云服务器后上传加解密主密钥 master S3 key, S3 密钥管理服务将其保存在服务器上, 其中主密钥保存的形式可能是明文也可能是密文。加密函数为 E , 主密钥记为 K 。

(2) 数据加密时, S3 密钥管理服务依据自身密钥生成算法产生加密配对密钥 k , 由云服务器对客户上传的需要加密存储数据 M 进行加密并存储到云存储服务器上, 数学表达式为:

$$E_k(M) = C_M \quad (1)$$

(3) S3 密钥管理服务用客户主密钥依据加密算法对密钥 k 进行加密并存储到云服务器上, 数学表达式为:

$$E_K(k) = C_k \quad (2)$$

(4) 数据解密时按照加密逆序过程进行, 解密函数 D 作用于 C_k, C_M 产生解密密钥 k 与数据 M , 数学表达式为:

$$D_K(C_k) = D_k(E_K(k)) = k \quad (3)$$

$$D_k(C_M) = D_k(E_k(M)) = M \quad (4)$$

事实上, 几大主流云存储服务商所使用的数据加密技术大同小异, 唯一的差别是当使用 Amazon S3 服务器端加密功能时, 密钥在 Amazon 手上, 而使用 Dropbox 时, 密钥在 Dropbox 手上。这两种方式实质上都一样, 虽然客户上传的信息在供应商的服务器上都是保持加密的, 但真正的信息拥有者客户却无法控制这些加密密钥。

3.2 Amazon S3 云存储服务的安全问题

事实上, S3 静态数据加密服务并不像看起来那么完美、安全, 其中一个重大的限制是, 这个解决方案没有提供外部密钥管理的功能, 主密钥跟 S3 认证相联结, 云端默认登录账户即合法账户, 如果云端账户认证被入侵, 黑客就

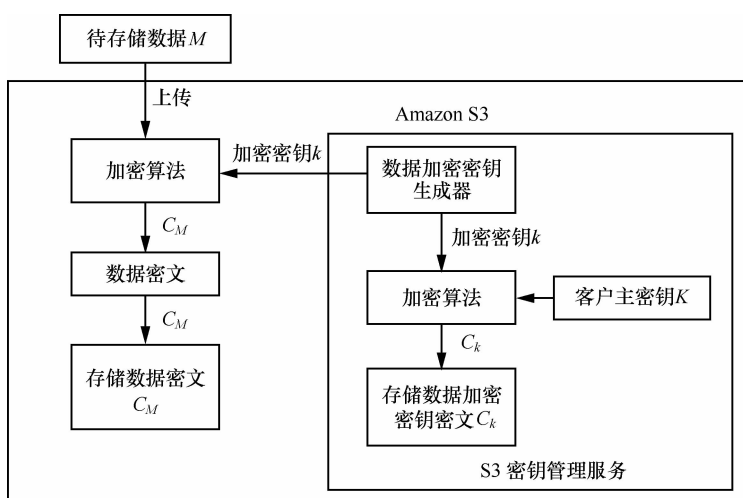


图 1 Amazon S3 工作原理示意

可以轻而易举地窃取客户存储在云服务端的一切数据。要黑掉一个云端账号并不用很高深的技术和太高的成本, 合理地运用社会工程学及一些简单木马病毒就可以很容易地获取云端账号及密码。

另一方面不能保证对 CSP 的绝对信任, 客户的所有数据都被 Amazon 掌控着, 包括密钥, 客户所谓的安全只能寄希望于服务端的绝对安全与服务商及管理人員的绝对良知, 显然事实并非如此, 特权管理人員的存在始终是无法逾越的安全隐患。事实上, 云存储的几个重要服务商如亚马逊、谷歌等都曾出现过各种安全问题, 并导致了严重的后果^[12,13]。因此, 寻找一种对于客户来说安全可靠的静态数据加密解决方案显得尤为重要。传统的云加密解决方案将用户加密密钥交付给了云服务商, 增加了泄露密钥的可能性, 要解决静态数据的安全问题, 必须使加密密钥受用户控制, 不存储在云中, 不暴露在风险中。

4 分离密钥云存储加密解决方案

从上面的分析中发现, 终端数据安全的压力主要集中在两个方面: 一是 Amazon S3 认证的安全问题, 二是主密钥的安全问题。针对以上问题, 本文设计了一套新的密钥管理方法分离密钥存储服务 (separated key S3), 试图解决上述两个问题。分离密钥存储服务的核心是虚拟密钥管理 (virtual key management, VKM) 服务。

4.1 虚拟密钥管理服务框架

在风险管理方面, 为了达到数据的安全, 虚拟密钥管理服务部署在云端服务器和存储服务器之间, 以确保云端服务器和存储服务器间的每字节数据都被加密了, 并且从



存储服务器移动到云服务器的每字节数据都只能通过客户来解密。虚拟密钥管理服务使用 VPD (virtual private data) 应用程序, VPD 是使用加密算法 (如 AES-256) 解密任何磁盘或存储阵列的虚拟设备。在数据进行存储和读取操作时, VPD 负责检索服务器配对密钥以及请求客户密钥, 完成存储数据的加解密。图 2 是其服务框架示意, 为简化叙述以下只介绍虚拟密钥管理服务实现数据加密的流程。

(1) 客户向云服务器递交数据加解密申请。

(2) 云服务器验证客户的加密申请后向虚拟密钥管理服务发出数据加密指令。

(3) 虚拟密钥管理服务在处理加密的过程中需要向客户请求主密钥。

(4) 客户通过对密钥请求的验证后返回主密钥, 并将需加密的数据传输给虚拟密钥管理服务。

(5) 虚拟密钥管理服务依据主密钥产生数据加密密钥对客户数据进行加密并配对分配存储到对应的云存储服务器上, 在此过程中主密钥和最终用于数据加密的密钥并不会以任何方式存放在服务器上, 这是保证数据安全的关键。

(6) 数据加密存储完成后云存储服务器向虚拟密钥服务器返回验证信息。

(7) 虚拟密钥管理服务向云服务器返回验证信息。

(8) 云服务器向客户返回验证信息。

图 2 中 (3)、(4) 这两个环节是数据加解密的关键步骤, 客户参与了每一次数据加解密的过程并起到了决定性作用, 只有当客户持有主密钥来参与才能完成上述加解密过程, 也就是说, 当客户的数据完成加密存储之后, 如需解密就必须提供主密钥, 否则, 数据就不能被解密, 于是客户成为唯一可以解密数据的人, 实现了将数据存储在云端而又保证

数据的完全私有化。

4.2 分离密钥存储服务存储加密解决方案

该方法主要受到银行保险箱模式的启发, 依据保险箱模式设置两个密钥, 一个交给客户保管, 将它称为主密钥 K , 另一个交给 VKM 服务, 由 VKM 依据客户数据及生成规则产生服务器配对密钥 k 。在数据进行加密之前, 虚拟密钥管理服务请求客户主密钥根据主密钥 K 与服务器配对密钥 k 依据本文所采用的安全算法 B , 产生新的数据加密密钥 K' 对存储数据进行加密存储, 其数学式表达为:

$$E(K, k) = K' \quad (5)$$

$$E_{K'}(M) = C_M \quad (6)$$

正如银行保险箱一样, 客户如果不提供服务商提供的服务器配对密钥, 就无法解密数据, 同样, 如果服务商不提供客户持有的主密钥, 也无法解密数据。其密钥关系如图 3 所示。

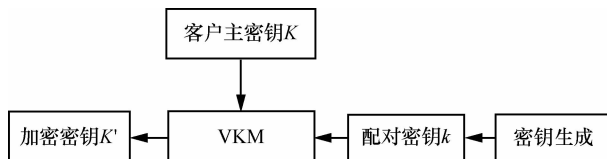


图 3 分离密钥存储服务密钥关系

除了将密钥分别交由客户和虚拟密钥管理服务保管外, 为了解决服务器配对密钥 k 的安全问题, 将服务器配对密钥通过客户的主密钥来加密, 而主密钥只有客户本人保管, 其数学式表达为:

$$E_k(k) = C_k \quad (7)$$

因此, 虽然上文所提到的特权管理人员可以通过入侵服务器持有配对密钥 k 的存储密文 C_k , 但却不能读取密钥, 因为它已经被加密了, 并且他所能入侵的程度仅限于

此, 因为遗留在服务器上的存储控制信息只有这么多, 在不知道客户主密钥 K 以及配对密钥 k' 的情况下, 特权管理人员无法获取最终的数据加密密钥 K' , 依据 AES 算法的安全性, 入侵客户数据的企图无法得逞。服务商尚不能破解客户存储在云端的数据, 那么黑客就更无能为力了, 就算黑客侵入云端服务器, 他所能窃取的也只是一堆毫无用处的密文, 存储在云端的数据 M 就不会被泄露, 数据实现了真正意义上的私有化,

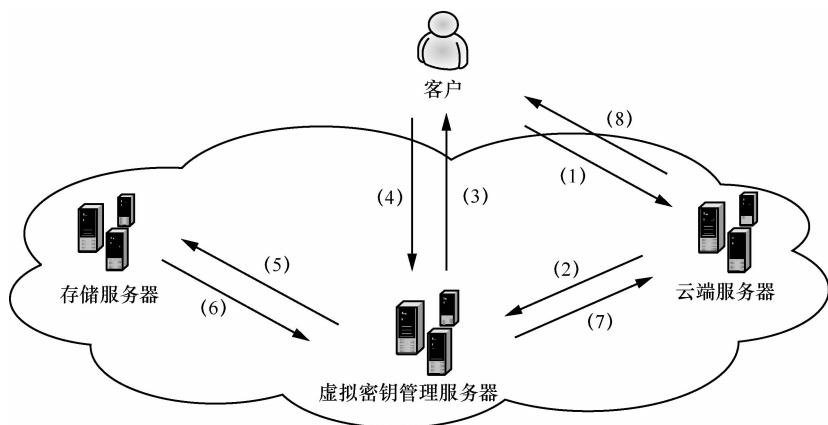


图 2 虚拟密钥管理服务框架示意

达到了数据安全的条件。此外,通过使用客户的主密钥来加密服务器配对密钥,完全缓解了终端数据保护的壓力。其工作原理如图4所示。

从上面的分析和对比中很容易就可以发现分离密钥加密存储方案的巨大优势,在该套方案中,云端认证和客户数据加解密分离,在数据加密过程中将密钥实现了分离,这样一来主密钥成为了保证数据安全的唯一关键,而客户成了世界上唯一知晓主密钥的人,主密钥可以被存放在只有客户本人才知道的地方,实现了真正意义上的云端静态数据私有化。

5 数据安全性

通过图2,可以分析出云环境中可能出现的两种隐私攻击^[14]。

- 内部攻击:由于数据存放在服务商提供的存储服务上,内部攻击者可以访问客户的数据。
- 外部攻击:数据在传递的过程中,外部攻击者可以通过窃听等方式窃取数据或者通过攻击服务器窃取资料。

本文讨论的方案主要针对第一种攻击。首先,破解加密数据的关键数据加密密钥 K' 并没有以任何方式存放在服务器上,当且仅当数据进行加解密操作时生成并使用,这是与传统的云存储方式相比最大的安全性改进;其次,假设攻击者突破了服务器的安全防护措施,攻击者最多只能窃取客户存储在存储服务器上的数据密文 C_M 和存储在

虚拟密钥服务器上的服务器配对密钥密文 C_k ,依据分离密钥加密算法,在不知道客户主密钥的情况下,攻击者无法解密服务器配对密钥 k ,也不能得到数据解密密钥 K' ,即无法解密加密后的数据密文 C_M ,客户存储在云端的私密数据不能被泄露,数据的安全性得到验证。对于第二种攻击,数据由静态成为动态,可使用安全的HTTPS加密传输协议保证私有数据和主密钥在传输过程中的安全。

此外,该方案的特点是能在很短的时间内创建一个安全的环境,可以将信任安全厂商或者云供应商的需求直接忽略掉。

6 性能测试

笔者采用已在校园内使用的Link1系统,Link1是为学校内用户提供服务的海量数据存储平台,由客户端和服务端组成,安装客户端后可以与服务端进行交互上传和下载数据。分别在服务器端物理主机上部署模拟Amazon S3及分离密钥存储服务,模拟实际Internet环境下的云存储服务。需要说明的是,为便于比较服务器参数,这里将分离密钥存储服务直接部署到物理主机上,而没有另外部署,这与设计方案略有出入。基于这种架构,进行了数据传输测试,测试环境:网络环境为100 Mbit/s以太网;服务器端为一台IBM System x3500 M4台式机,CPU为Xeon E5 4核2.4 GHz,内存8 GB,硬盘为SATA 300 GB(7 200转),网卡为1 000 Mbit/s以太网;客户端为一台华硕S400E3217CA,CPU为Intel酷睿3双核1.8 GHz,内存4 GB,硬盘为SATA 500 GB(5 400转),网卡为1 000 Mbit/s以

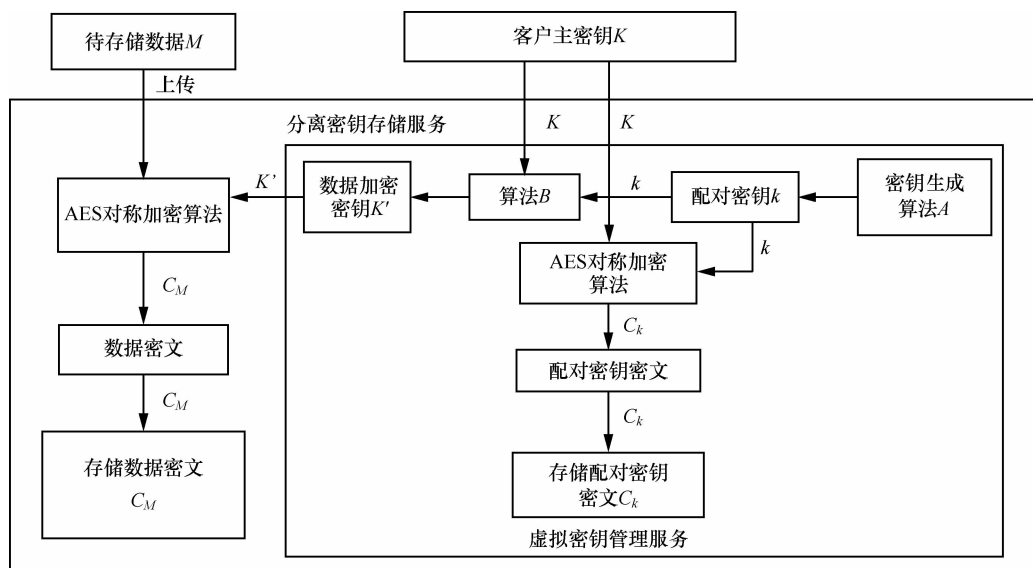


图4 分离密钥存储服务工作原理



网卡。

笔者使用了一个 1 GB 的文件进行大文件传输测试,其中“Origin”代表现有的存储环境,没有加密功能状态下的性能;“Amazon”代表部署了具有普通数据加密功能,并不能保证数据私密性模拟 Amazon S3 后的性能;“Separated”代表部署了可以保证数据的安全性和私密性的分离密钥存储服务后的性能。表 1、图 5 为测试结果。

表 1 大文件传输时服务器端性能数据

类型	服务器 CPU 占用率	下载速度 (Mbyte/s)	服务器 CPU 占用率/下载速度
Origin	3.23%	11.2341	0.28752%
Amazon	14.76%	3.8521	3.83168%
Separated	18.45%	3.8468	4.79619%

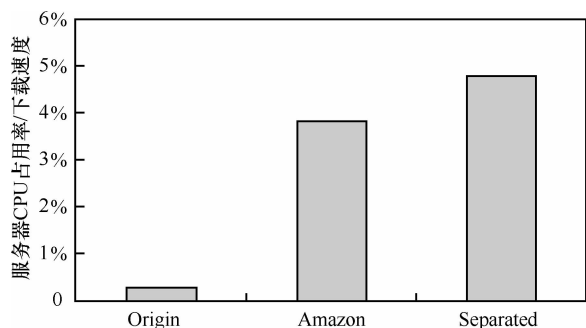


图 5 大文件传输时服务器端 CPU 占用率/下载速度

表 1 和图 5 表现了服务器端在原始、加入 Amazon S3 和加入分离密钥存储服务这 3 种状态下 CPU 占用率与下载速度之比,在 3 种状态下,除回复主密钥外客户端并未有附加运行程序,所以 CPU 占用率和下载速度并没有明显变化。从表 1 和图 5 中可以看到,在服务器端和原始的情况相比,Amazon S3 的部署引入了 13 倍多的性能损失,而分离密钥存储服务则引入了 3 倍多的性能损失。

从性能测试的情况来看,分离密钥存储服务引入的性能损失远小于使用 Amazon S3 引入的性能损失,为了保护客户数据的私密性,实现数据的私有化,这样的损失对于应用来说是可以接受的。

7 结束语

云存储是一种趋势,但是数据的安全性、可靠性、隐私保护等问题是这项业务能够得到广泛应用的决定性因素。原有的云存储模式所存在的弊端及问题已经困扰行业多年,在静态数据加密存储方面,只有实现真正意义上的数据

私有化才能保证数据的安全,保证数据拥有者的利益。本文针对云端静态数据安全所设计的云存储解决方案,与客户端加密模式相比,使客户无需再管理复杂的密钥,也无需在性能不高的客户端进行复杂而耗时的加解密操作,克服使用一般方法带来的弊端,实现了数据的真正意义上的私有化,保证了云存储的数据安全,且具有较高的应用价值。

参考文献

- 林秦颖,桂小林,史德琴等.面向云存储的安全存储策略研究.计算机研究与发展,2011,48(z1):240~243
- 黄永峰,张久龄,李星.云存储应用中的加密存储及其检索技术.中兴通讯技术,2010,16(4):33~35
- 杨洋.存储层:动态数据、静态数据分而制之.中国计算机报,2008-09-22
- 孙磊,戴紫珊,郭锦娣.云计算密钥管理框架研究.电信科学,2010,26(9):70~73
- Wang G J, Liu Q, Wu J. Hierarchical attribute-based encryption for fine-grained access control in cloud computing. Proceedings of the 17th ACM Conference on Computer and Communications Security, Chicago, IL, USA, 2010
- De Capitani di Vimercati S, Foresti S, Jajodia S, et al. Encryption-based policy enforcement for cloud storage. Proceedings of IEEE the 30th International Conference on Distributed Computing Systems Workshops, Genova, Italy, 2010:42~51
- 刘帆,杨明.一种用于云存储的密文策略属性基加密方案.计算机应用研究,2012,29(4):1452~1456
- 密钥管理:云计算系统保护数据的核心机制.http://www.ciotimes.com/cloud/caq/52436.html, 2011
- 王天予.同态加密用于云端数据泄露或成往事.http://www.enet.com.cn/article/2011/0809/A20110809895739.shtml, 2011
- Amazon simple storage service (Amazon S3).http://aws.amazon.com/s3/
- 吴吉义,傅建庆,张明西.云数据管理研究综述.电信科学,2010,26(5):34~39
- 冯登国,张敏,张妍.云计算安全研究.软件学报,2011,22(1):71~83
- 李战宝,张文贵.云计算及其安全性研究.信息安全,2011(1):5~7
- 黄汝维,桂小林,余思等.支持隐私保护的云存储框架设计.西安交通大学学报,2011,45(10):1~6

【作者简介】王志中,男,解放军重庆通信学院硕士研究生,主要研究方向为计算机应用技术、军事信息安全。周城,男,重庆市信息安全重点实验室副教授、硕士生导师,主要研究方向为计算机应用技术、军事信息安全。牟宇飞,男,解放军重庆通信学院硕士研究生,主要研究方向为数字图像处理。

(收稿日期:2012-11-05)