2012 年 3 月 第 3 期
Vol. 33 No. 3 2012

云计算安全问题研究综述

杨 健[1,2], 汪海航[1], 王 剑[1], 俞定国[1]

[1](同济大学 电子与信息工程学院,上海 201804)
[2](大理学院 数学与计算机学院,云南 大理 671003)
E-mail:sbjc1215@126.com

摘 要: 随着云计算的蓬勃发展,越来越多的企业和个人将他们的存储和计算需求付诸于云端.然而云计算的安全仍不容忽视,是当前的一个研究热点.对近年来云计算安全相关的研究成果进行总结,主要集中于数据安全,身份认证以及访问控制策略方面.也介绍了与可信计算技术相结合的云计算安全的相关研究框架和项目.根据这些研究成果,认为将可信计算与云计算思想相结合,建立"可信云计算"是未来云计算安全研究的一个重要方向.并且在最后提出了"可信云计算"发展的几个可能的研究主题.

关 键 词: 云计算;数据完整性;身份认证;访问控制策略;可信云计算
中图分类号: TP393 **文献标识码**:A **文 章 编 号**:1000-1220(2012)03-0472-08

Survey on Some Security Issues of Cloud Computing

YANG Jian[1,2], WANG Hai-hang[1], WANG Jian[1], YU Ding-guo[1]

[1](*Department of Electronic and Information Engineering, Tongii University, Shanghai* 201804, *China*)
[2](*Department of Mathematics and Computer Science, Dali University, Dali* 671003, *China*)

Abstract: More and more organizations and individuals outsource their storage and computing needs into a new economic and computing model, which is commonly referred to as cloud computing. As one of the hottest research issues, cloud computing security is also concerned by many scholars. This paper survey the recent advances in some security issues of cloud computing, which mainly focuses on the topics of data security, identity authentication and access control policy in cloud computing environment. It also introduces some enhancement frameworks and projects about the cloud security integrated with the Trusted Computing. According to these surveys, this paper believes that the combination between trusted computing and cloud computing will be a promising direction of the future cloud security researches, and proposes some interesting research issues in the end.

Key words: cloud computing; data integrity; identity authentication; access control policy; trusted cloud computing

1 引 言

云计算从其概念提出一开始,就受到了广泛的关注,许多国际厂商分别推出自己的云计算解决方案,典型的有 Amazon 的弹性计算云 EC2 和简单存储服务 S3,IBM 的 Blue Cloud, Google 的 App Engine 等.与此同时,云的各种安全问题也逐渐走向世人面前,成为目前阻碍云计算发展的最大阻力.美国 Gartner 公司总结了七条云计算安全风险[1],它们分别是:

1) 特权用户访问风险;
2) 法规遵守风险;
3) 数据位置不确定风险;
4) 共享存储数据风险;
5) 数据恢复风险;
6) 调查支持(数据跟踪功能)风险;
7) 长期发展风险.

而云安全联盟 CSA 所提供的安全指南 V2.1[2]从 13 个方面对云计算中的安全主题进行导向性的论述,并提出了相应的建议.其中涉及到的云计算中特有的安全问题概括起来主要有:

1) 用户数据存放在外部的数据中心,需要增加加密措施保证数据安全,并采用一定的认证和访问控制策略;

2) 为了保证数据可恢复性,通常采用冗余存储的手段,这需要特定方法保证多个版本数据的一致性和完整性,并采用特定的方法进行审计(audit);

3) 应用以网络上松耦合的云服务形式存在,需要在 5 个方面进行安全增强(例如应用安全结构、软件生命周期等);

4) 加密机制和密钥管理机制的改变;

5) 虚拟化是云的 3 个参考模型 IaaS/PaaS/SaaS 的重要的理论基础,而虚拟化同时带来安全问题和虚拟机的安全和管理问题.

收稿日期:2010-10-18 收修改稿日期:2011-01-17 基金项目:上海市科委科技支撑计划项目(072712036)资助. 作者简介:杨 健,男,1976 年生,博士研究生,讲师,CCF 学生会员,主要研究方向为云计算,网络安全,智能信息系统;汪海航,男,1965 年生,教授,博士生导师,主要研究方向为信息安全、智能信息系统;王 剑,女,1978 年生,博士研究生,讲师,主要研究方向为网络安全、可信计算;俞定国,男,1976 年生,博士研究生,主要研究方向为信息系统安全.

可以看出,云计算的主要特性有按需服务,效用付费,网络共享的存储和计算资源池,迅速弹性化部署和泛网络访问,而这几个特性也导致云的安全问题与传统网络应用面临的安全问题有所不同.这些安全问题受到了广大研究者的广泛关注,本文就目前的部分领域研究现状及成果进行综述,并提出一些建议性的发展思路.主要论及的研究方向有:

1)数据及身份的保密性、完整性保护;

2)用户身份及操作的隐私保护;

3)审计和数据取证.

下面的章节按照如下的结构组织:第二节介绍数据的保密和安全问题;第三节介绍云数据和身份的隐私保护问题;第四节介绍数据的审计和数字取证问题;第五节介绍一些综合研究成果;第六节介绍可信云计算并总结和展望.

2　云中数据的保密和安全问题

云计算中,用户将数据存储在云端,因而不再拥有对自己数据的完全控制能力,要求云服务商(Cloud Service Provider, CSP)提供有效的安全保障,使其能够信任新环境下的数据安全及完整性.相比于传统计算,这种数据新的访问和控制模式带来了新的安全挑战.

2.1　数据安全新问题和新方法

就云计算的3种参考模型来说,IaaS一般以Web服务的接口形式提供,SaaS服务常通过Web浏览器访问,PaaS服务则用上述两种技术的结合来实现.而网络中应用层协议传输数据和参数主要以XML为载体.文献[3]介绍了云计算中涉及Web服务和浏览器的一些安全问题.认为仍存在针对XML签名的有效攻击,并且浏览器安全问题不但需要使用传输层安全技术加以解决,还应当在浏览器核心代码中加入XML的加密机制.由于目前浏览器存在的安全问题,使得基于浏览器的认证也存在漏洞(联邦身份管理需要浏览器存储安全令牌,而浏览器并没有对XML加密的机制).此外,对云服务完整性和使用虚拟机的特点,也存在恶意软件注入、元数据欺骗及针对服务器的Dos攻击.因此,从应用的角度想要提高云计算的安全性,就必须从Web浏览器和Web服务框架两个方面来增强安全能力.

针对Web 2.0应用,文献[4]提出了一个安全文件存储服务的文件系统框架,利用目前安全的客户端跨域(Client Cross-domain)通信机制的研究成果,给Web服务提供一个独立的文件系统服务,将用户数据的控制权返还给用户,提高了数据的可控性,降低了应用服务器管理用户数据的访问控制策略的压力.文献[5]也提供了一个基于云计算的安全文档服务机制.通过将文档的内容与格式相分离,并在传输到外部之前对内容进行加密的方式,降低内容泄漏的风险.除此之外还包含一个优化的文档授权访问方法.

一个加密的网络文件系统要在支持随机访问的基础上保证文件的机密性和完整性.现在流行的设计是把Merkle hash树[6](一种使用hash函数的树形结构选择性泄漏协议)和加密的块密码结合.文献[7]提出一个新的基于MAC树(一种对每个数据块都进行密钥认证的方法[8,9])的加密网络文件系统结构,相比于Merkle树的结构有更好的性能,以更低的成本提供了完整性保护.

数据存储在云中是以分布式文件系统的形式存在.如何在允许数据块动态操作的基础上验证数据正确性并定位错误数据是一个需要解决的问题.文献[10]利用了分布式文件系统纠删码的研究成果,用户预先计算数据块的验证令牌,而服务器在接收到用户验证挑战后,根据挑战生成指定块的"签名"并返回给用户.用户通过比较这些签名与预计算的令牌来判断数据正确性.用户端和服务器端使用的通用散列函数具有同态保持的属性.这种方法实现了存储数据正确性和数据错误定位的功能,并同时支持安全高效的数据块动态操作:数据的更新、添加和删除.

上述的这些方法在云计算环境中有效提高数据块保密性和完整性,减少了数据块完整性认证时证书所需的存储空间,是未来云计算数据完整性验证研究的重要方向.除了使用传统方法以外,一种趋势就是利用用户的数字身份完成数据加密和身份认证功能:

云是一个存储资源和计算资源开放共享的环境.云服务一般通过给用户提供数字身份标识用户,同一用户不同服务需要管理不同的加密和签名信息.在有多个公有和私有云组成的混合云中,用户身份及对应的密钥难以管理.而基于身份的加密和签名系统(IBC)[11]恰能弥补这种不便.IBC中,使用用户身份相关信息作为公钥,而用户的私钥由一个公开可信的PKG(Private Key Generator)结合用户身份来生成并安全传输给用户.然而,这会导致一个严重问题,即密钥集中于PKG的管理问题.并且系统的可扩展性也不高.针对这个问题,层次型的IBC被提出来[12,13]以提高传统IBC的可扩展性.文献[14]中,提出一种联邦身份管理机制.其主要思想是:在所有云之上有一个权威PKG,每个子域云(公有云或私有云)也有自己的PKG,子域中的用户和服务器由本域的PKG管理身份密钥,而权威PKG负责给子域云分配ID,从而形成三层HIBC(如图1所示).这种结构简化了云中密钥分配及相互认证,并且缓解了PKG的密钥托管问题(只有本地的PKG知道用户密钥).

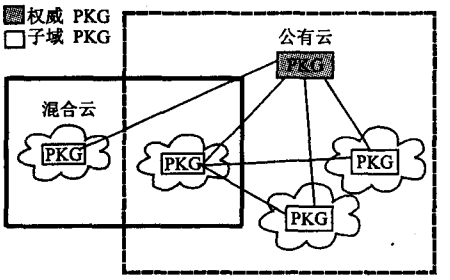

图1　层次型基于身份的密码系统

Fig. 1　PKGs in hierarchical identity-based cryptosystems

2.2　身份认证及访问控制策略

用户要使用云存储和计算服务,必须要经过云服务商CSP的认证,而且要采用一定的访问控制策略来控制对数据和服务的访问.各级提供商之间也需要相互的认证.因此云计

算中的认证和访问控制是一个重要的安全领域研究问题.

安全套接字层认证协议(SAP)曾用于云计算中的认证,但该协议复杂且计算和通信的负载较大.而在云计算中,用户都有自己的数字 ID,因此一个趋势是用身份 ID 作为认证的基础.文献[15]在基于身份的加密(IBE)和签名(IBS)的基础上,提出一个用于云计算和云服务的加密和签名的基于身份的认证协议(IBACC).相比于 SAP,它不需要认证证书,认证协议很好地满足了云计算的需求,通过在仿真平台 GridSim 上的实验表明性能上比 SAP 更具有优势,且在用户端负载低.另外,[14]提出的联邦身份管理(Federated Identity Management,FIM)在层次型基于身份的加密和签名的基础上也实现了基于身份的认证功能.

云计算本质是一个分布式的系统,因此各个节点(包括各个服务)之间的访问控制策略的互理解能力也成为云计算安全领域中的一个重要问题.Web 服务中的 WS-Security 等规范和语义 Web 技术为异质的语义互操作提供了解决.文献[16]提出一个新的语义访问控制策略语言(SACPL),并设计了面向访问控制的本体系统(ACOOS)作为 SACPL 的语义基础.它能够有效解决分布式访问控制策略之间的互操作问题,扩展了语义 web 在安全领域的研究,提供了云服务之间认证的一个语言描述环境.

在访问控制策略方面,文献[17]基于数据属性来定义和增强访问控制策略,其理论基础包含基于属性加密的密钥策略(KP-ABE),代理重加密(Proxy Re-encryption,PRE)和惰性重加密(Lazy Re-encryption,LRE)三个方面.KP-ABE[18]是一个利用了双线性映射和离散对数问题的一对多的公钥加密通信机制.允许单个数据拥有者与多个数据使用者进行安全数据分发;而 PRE[19]是一种加密机制,其中的半信任代理能够将 Alice 公钥加密的密文转换为另外一份密文,使得在不查看原始明文的基础上,密文能够被 Bob 的私钥解密,如图 2 所示.[17]中用一个属性集与每个数据文件关联,并给每个用

图 2 PRE 的基础结构

Fig. 2 Proxy re-encryption

户指定一个定义在这些属性集上的的访问结构.用 KP-ABE 来管理数据拥有者与数据使用者之间的信息交换密钥.但这会导致数据拥有者异常繁重的计算任务(尤其是在移除用户时,数据拥有者必须自己计算更新所有与移除用户关联的文件的密钥).因此,结合 PRE,将繁重的密钥计算任务委托给云服务器,而不需要向云服务器揭示底层的文件内容.这样的结构降低了用户端的计算负载,并保证了数据的安全.为了进一步降低云服务器的计算压力,利用了 LRE,允许云服务器积累多个系统操作的计算任务进行批量计算.云服务器的计算复杂度与系统属性个数成正比,且与用户访问结构树的大小成线性关系,与云系统中的用户个数无关,从而获得可扩展性,也防止用户隐私信息在云服务器端泄漏.这种访问控制方法在实现细粒度访问控制的基础上,同时保证了可扩展性、数据安全和用户隐私.

云中的数据存储和使用方式多种多样,其中有一部分具有"拥有者写使用者读"的特性.文献[20]针对这种类型的数据存储和访问机制提出了一种访问控制方法.使用不同的对称密钥加密每一个数据块,并采用密钥导出方法[21,22]以减少数据拥有者和终端用户需要维护的秘密数量.密钥导出的基本思想是通过一个层次形结构生成数据块加密密钥.层次中的每个密钥能够通过结合其父节点和一些公有信息使用单向函数导出.这种结构类似于前述的 Merkle hash 树.另外,文中还采用了服务器的过加密(Over Encryption)[23]以获得终端用户之间的数据隔离特性;采用惰性移除[24]来阻止已被移除的用户继续访问更新的数据.文中还设计了方法来解决数据更新和用户访问权限变化的问题.从计算、存储和通信开销上与其他现存类似的解决方案比较,该方法在"拥有者写使用者读"这种特定的云应用环境取得了更好的可扩展性和安全性.

2.3 虚拟机安全和自动化管理

虚拟化和虚拟机技术是云计算概念的一个基础组成部分.在 SaaS 的模式中,应用构建在虚拟化平台之上,用户以透明的方式与其他用户共享物理计算资源来运行服务;在 IaaS 和 PaaS 的模式中,应用以虚拟机或虚拟化平台的模式提供给用户使用.除了传统的网络、系统和软件安全外,首先要求虚拟机在共享物理计算资源和存储资源时具有隔离性,另外要求虚拟机监控程序是可信的,并且不能涉及用户隐私相关信息,而且虚拟机本身要安全、可信,具有内省机制.国内外学者在这些领域也开展了相关的研究.

很多情况下,CSP 并不是虚拟机映像的提供商,因此,必须有较好的方法来对虚拟机映像进行管理.VMware 的 Virtual Appliance Market Place 和亚马逊的 EC2 提供了映像库的概念,然而只是简单的存储和提取.文献[25]提出了一个映像管理系统 Mirage,控制对映像的访问,跟踪映像的来源,给用户和云管理员提供有效的映像过滤和扫描机制,检测和修复映像漏洞.然而该系统仍有很多需要完善的地方,例如:没有提供自动检测和过滤用户隐私的问题,恶意软件的扫描不能完全保证映像中不存在恶意的软件等等.

针对虚拟机与监控器及物理资源之间存在相互的配置需求,文献[26]提出了虚拟机契约(VMC)的概念.通过对开放虚拟机格式(OVF)标准的扩展可以对 VMC 进行表达,并以统一的方式管理虚拟机.其中,OVF 是得到 VMWare 等许多大厂商支持的一个工业化标准,一个 OVF 包包含了一个 XML 格式的 OVF 描述符,用以指定虚拟设备元数据配置,还包括了一个虚拟磁盘文件的集合.VMC 为虚拟机在大型数据中心和云计算环境中实现自动控制和管理提供了一个思路.除此之外,在虚拟机检测,虚拟网络访问控制和灾难恢复等方面也有一定的辅助作用.

现有的对虚拟化安全的研究大多基于两个假设:

1)虚拟机监控器具有被虚拟化的软件(运行于虚拟机上的客户 OS 等)的先验知识;

2)现存技术需要从客户操作系统启动之时起监控客户

的虚拟机.

第一个假设存在一个语义鸿沟(基于虚拟 OS 的符号表,而不管运行时的虚拟机内存分配是否与符号表匹配),而第二个假设在云计算复杂的虚拟化应用环境下显然不适合,因为虚拟机除了运行状态外还存在其他可以被恶意攻击的状态.既然从虚拟机外部不能对虚拟机内部的运行进行监控,就要求虚拟机具有自我检查的机制.文献[27]在只假设硬件状态完整性的条件下,从已知硬件元素(例如中断描述符表 IDT)开始,自动探查运行的 VM 的代码,评估它以及它依赖的数据结构的完整性.

客户 OS 的内核完整性验证包含如下 4 个步骤:

1)从虚拟 CPU 中读取 IDT 的位置;

2)使用 IDT 内容的分析、内存代码块的 hash 值以及已知操作系统的白名单,来确定 VM 上运行的客户 OS 的完整性;

3) 使用运行 OS 的信息,利用恰当算法发现其他操作系统链接的结构(如系统调用表、进程列表和加载的内核模块等);

4)使用适合的白名单继续分析已发现的数据结构,进行完整性检验.这个方法允许在虚拟机生命周期的任何阶段开始对其完整性进行验证.可以看到白名单在该方法中的重要性.在假设攻击者不能破坏虚拟机硬件的基础上,这种 VM 的内部检查方法是有效的.

云计算的一种典型的组织结构是,服务提供商和基础设施提供商是分离的.服务提供商需要在基础设施提供商的设备上部署自己的应用.这就要求各个不同的应用之间具有有效的隔离措施.另外,多个虚拟机部署在同一个物理机器上时,也需要有效地隔离各个虚拟机对物理资源的使用.多核处理器的末级高速缓存(LLC)可以被不同的虚拟服务共享,这就造成了资源隔离的困难.文献[28]针对这个问题,利用缓存层次感知的内核分配技术和基于页面着色的缓存隔离技术.前者是根据缓存的组织结构对内核进行分组,共享 LLC 的内核在同一组里,从而可以匹配服务的 SLA 资源需求以及达到内核级别的隔离.页面着色是一种用来确保虚拟内存中连续页的访问能够最好地利用处理器缓存的性能优化技术.这使得将安全及性能隔离的约束加入到云计算服务等级协定(SLA)成为可能.也有助于在虚拟机级别进一步提高 QoS 和安全保障.

2.4 新的安全问题

针对云计算的特性,各种安全方案相继被提出,但也逐渐引入了其他一些安全问题.例如在隐私保护问题上,流行的隐私增强技术如 OpenSSL、OpenVPN 等,通过建立一个加密通道,隐藏传输的内容和请求的 web 站点地址.文献[29]针对这种隐私增强技术提出了一种攻击方法,在可观察 IP 包大小的归一化频率分布上应用普通的文本挖掘技术,建立多项式纯贝叶斯分类器.实验中正确检测出了 97% 的 web 站点请求.这是一种利用数据挖掘技术的流量分析攻击方法,很多流行的隐私增强技术在这个通用的指纹攻击面前是脆弱的.

3 云数据隐私保护问题

在云计算环境中,一个最重要的特征就是用户数据不再存放于本地,而是存放到云端,其中的敏感数据会带来隐私保护问题.虽然很多云安全指南建议人们不要将敏感数据放到云端,然而这并不是长久的解决之道,而且会抵消云计算带来的好处,阻碍云计算的进一步发展.因此,采用何种方法保护用户隐私,成为当今研究的一个热点.另外,数据存放到云端,用户在利用云服务使用数据时,需要根据使用来付费,而且部分的本地地方法律以及商业运营(如在线广告)也对数据的存放和使用有一定的需求,这就需要有效的机制在不泄漏敏感数据内容的基础上,对数据的存放和使用进行监控和审核.

3.1 隐私管理

大多数云计算中的隐私管理强调云服务器的作用,主要利用云端的管理组件实现隐私管理.然而,文献[30]中描述了一种基于用户的隐私管理器.提供了一种用户为中心的信任模型,在服务提供商能够协作的假设下,帮助用户控制他们的敏感信息.并且,通过使用混淆(Obfuscation),在即使没有服务商的协同工作,甚至服务商是恶意的情况下来保护数据隐私.文献[31]描述了一个在云计算环境下的隐私管理器.在该体系结构中,用户私有数据以加密形式通过隐私管理器提供给云.基于一个用户和隐私管理器共有的密钥,隐私管理器对数据进行混淆和解混(De-obfuscation),以便在云端隐藏数据真实内容,在客户端给用户显示真实结果.并且,隐私管理器充分利用了 TPM 来保护混淆密钥,进一步增强了隐私保护特性.文中介绍了混淆的代数描述和几个特定应用场景的简化描述,并对SQL查询语言提出了具体的混淆实例和结果.

上述两个隐私管理器都应用了混淆的技术.混淆是用户对私密数据 x 进行某些函数 f 求值 f(x),并将 f(x)上传至服务器.服务提供商在不知晓 x 的情况下,针对某项云服务,对 f(x)求 f′(x),并将 f′(x)作为服务结果返回给用户,用户再进行进一步的处理.虽然混淆是一个保护用户隐私的好方法,但仍然有需要改进的地方.例如,混淆的过程通常在用户端完成,这就要求用户有一定的计算能力,在频繁进行计算的时候会造成计算瓶颈;另外,尽管存在某些特定的运算可以在不需要揭示实际数据的情况下得到一致的结果,但仍有大量运算在没有明确输入的情况下得不到正确的结果.这方面的研究,如加密数据查询,也得到越来越多学者的关注.

用户数据存放在云端,一方面要求云服务商能够根据他们的查询提供正确的查询结果,另一方面又不希望云服务商知晓用户数据的实际内容,也即希望在加密的数据上实现数据查询等计算功能.文献[32]提出了一种隐私保护的关键字查询方法.利用了一种带有关键字查询的公钥加密方法(PEKS):在 bob 与 Alice 传递邮件的场景中,利用 Alice 提供的一个陷门(trapdoor),使得第三方代理在不知道邮件内容的情况下测试某个单词是否包含在 Bob 发给 Alice 的邮件中.该方法允许服务提供商部分参与内容解密并进行相关内容的查询,但不能由此得到全部明文,这可以在隐私保持的条件下减少用户端信息处理(加密/解密)的压力.文中还通过改进 PEKS 提出一个有效的隐私保护关键字搜索机制.从语义角度证明了该方法是安全的.

传统的在加密数据上的关键字查询只能针对确切的关键字. 文献[33]提出在加密云数据上进行有效的模糊(fuzzy)关键字查询,并保持关键字的私密性. 当用户的查询与以前定义的关键词完全匹配时,就返回匹配的文件,否则,基于关键词相似度语义返回最可能匹配的文件. 文中利用编辑距离(edit distance)量化关键词相似度,并使用通配符描述相同位置的编辑操作(插入/删除/修改一个字母),从而构建出高效的模糊关键字集合. 在这个表示的基础上,进一步提出了构建数据文件索引,查询相关文件的方法. 分析表明这是一个安全的高效的模糊关键字查询方案.

加密数据的隐私保护查询查询还有一种特定应用环境就是,客户搜索云服务器的信息,而不希望服务器知道他的 ID 或查询的内容;当然,从云服务本身来说,也不希望这种查询获知与查询无关的其他信息. 这种特定的应用可以看作是一种强隐私保护场景. 例如为了调查取证,警察查询某个人的银行账户的特定信息. 这种特定应用可以称之为安全匿名查询(Secure Anonymous Database Search, SADS). 这种特定情况与前述隐私保持的区别在于,数据查询者的 ID 也需要对服务器保密,与此同时,数据拥有者还必须阻止数据的非法使用. 文献[34]针对这种情况提出了一种解决方案. 该协议有两个中间的代理实体:查询路由器(query router, QR)和索引服务器(Index Server, IS). IS 存储了数据拥有者构建的加密查询结构,并且在不知道查询内容和任何底层数据库的条件下执行提交的查询. 而 QR 连接查询者和 IS,且不会将任何查询者的 ID 暴露给其他实体. 为了获得查询的效率,使用 Bloom filter[35] 作为查询结构. 为了保护查询者 ID 不被服务器探知,还定义了一种可重寻路加密(Re-routable Encryption)协议,这种机制类似于代理加密[36] 和通用重加密(Universal Re-encryption),即允许一个不可信代理转换 A 加密的密文,使之能被 B 所使用,并且代理不会知道任何关于明文及 A/B 的密钥的信息.

隐私保持的一个具体应用就是电子医疗记录系统中病人的隐私保持问题. 文献[37]中对这个问题进行了探讨,提出应当经由加密和访问控制来增强保护. 认为应当使得病人生成和存储信息的加密密钥以便于隐私保护,并形式化地描述了病人控制加密(PCE)方案的需求. 针对客户端加密对性能影响的疑虑,建立了一个高效的系统,允许用户与他人共享部分访问权利,并且在相应的记录上进行搜索.

可以看到,云的特殊存储结构使得隐私保持成为一个关键的安全问题. 目前应用最多的方法就是对上传到云端的数据进行混淆和加密. 同时,还应当有有效的查询和用户验证机制,在云服务器不能获知具体数据内容的条件下,获得云服务的数据处理结果. 并且,这种隐私保持机制应当是用户可控的.

3.2 从应用设计角度考虑隐私相关的设计原则

从应用设计的角度考虑隐私相关的设计原则,文献[38]认为在云计算软件开发的各个周期都应当考虑隐私保护问题. 文中提出了几个面向隐私保护的设计原则:

1)进行隐私影响评估;

2)在系统的不同设计阶段评估隐私;

3)在适当的时候使用隐私增强技术(PETs);

4)云系统设计者,构建者,开发者和测试者应遵循以下 6 个隐私实践建议:发送和存储最小化的个人信息、保护个人信息、最大程度实现个人控制、允许用户选择配置隐私管理、规定和限制使用数据的目的、提供反馈.

4 数据取证及审计问题

云计算中用户数据不再被用户本地拥有,因此需要有方法让用户确信他们的数据被正确的存储和处理,即进行完整性验证;另外,从涉及到数据安全和使用的法律和网络监管角度,也需要一种机制能够远程、公开地对数据进行审计. 并且,这种审计必须以不泄漏用户隐私信息为前提.

已有一些方法提供对远程数据的验证. 例如[39,40]利用基于 RSA 的同态标签(Homomorphic Tag)实现"可验证数据持有"(Provable Data Possession)模型. [41]在此类研究的基础上进行了扩展,通过使用用于块标记认证的经典的 Merkle hash 树,改进了可恢复证据(Proof of Retrievability)模型[42,43],在保证不影响数据块的插入、修改和删除等动态操作的基础上,实现了利用第三方审计(TPA)完成隐私保持的数据完整性验证. 这种方法不需要云用户的实时参与,且避免用户隐私的泄漏. 文献[44]也利用 TPA 实现外部数据的安全、高效审计. 它利用带有随机伪装(Random Masking)的同型验证器(Homomorphic Authenticator)以实现隐私保持的公开数据审计,并进一步探索了双线性聚集签名技术以扩展到多用户环境,使得 TPA 能够同时执行多个审计任务. 在可恢复证据(PoRs)方面,文献[45]提出一个用于设计 PoRs 的理论框架,改进了以前提出的 POR(Juels-kaliski 协议)结构,在其基础上提出一个新的变体结构. 它支持抵御完整的 Byzantine 攻击模型.

文献[46]建立了一个可扩展的运行时完整性验证框架 RunTest,以保证在云基础设施上数据流处理结果的完整性,并当检测出不一致结果时,能明确定位恶意服务提供商.

记录了数据对象的拥有关系和处理历史的安全数据起源(provenance)对于云计算中成功地进行数据取证是很重要的. 文献[47]基于双线性配对技术提出一个新的安全数据起源追踪方案. 该方法为在云中存储的敏感文档提供了保密性,对用户访问提供匿名验证,对问题文档进行起源跟踪.

5 其他一些安全研究思路

在我们利用 Internet 共享云计算的按需提供的存储和计算能力带来的优势的时候,遗产应用和受限设备应用(例如移动终端)却没有很好地利用这种模式. 文献[48]针对这个现象提出了一种解决方案. 其目标是建立弹性应用,将各种资源受限平台与云中弹性化的计算资源联接起来. 一个弹性应用包含一个或多个 weblets,每个可以在受限设备或云中运行,并且可以根据计算环境和设备上用户喜好的动态变化,进行迁移. 首先,提出了在设备端和云端运行的 weblets 之间认

证和安全会话管理的解决方案,然后提供了安全迁移 weblets 以及如何授权云端 weblets 访问敏感用户数据,例如经由外部的 Web Services. 这种解决方案的一些规则能应用于其他一些云计算场景,例如企业环境下私有云和公有云的应用迁移.

文献[49]提出了将云计算和网络安全结合起来进行"云"防火墙技术的研究思路. 其主要思想是利用云的特性将防火墙的被动的保护转换为动态、协作和主动的保护:一个地方受到攻击,某些终端就立刻告知其他地方以阻止攻击的进一步蔓延并采取相应的抵御措施. 其最重要的特性是动态性和智能型,它充分利用了云以动态地、实时地采样和共享威胁信息,来实现实时的、主动应对的安全服务. 然而,云防火墙的发展仍处于孕育阶段. 一种实现趋势是:通过 Internet 将防火墙软件平台与客户连接起来形成一个巨型的木马/恶意软件监控器,每个用户对云安全做出贡献并与其他用户分享安全信息,类似于多个防病毒软件商(例如瑞星)提出的云安全概念;另一种趋势是:在全世界建立足够数量的服务器来搜集应用请求,通过在云顶(thetopofcloud)来判断这些请求的安全性.

对于终端用户使用浏览器访问云服务的体验来说,文献[50]认为目前基于 SSL 的用户接口并不能很好地让用户更加信任云计算技术,并且显得复杂、难以理解. 研究者探索在最广泛部署的浏览器(IE7)上关于 SSL 证书的接口扩展,提供了一个接口会话的可选集合,通过一个有 40 个参与者参加的用户研究比较扩展的用户接口的效果. 可选的用户接口集合提高了使用者的信任度,易于找到信息,易于理解. 这为目前浏览器用户接口的改进提供了余地. 文章同时建议各大浏览器厂商从用户接口上对浏览器进行改进,以增强用户的安全体验.

6　可信云计算

随着云计算进一步的发展和壮大,各种安全问题逐渐被认识和发现,各种解决方案也陆续被提出. 然而,在复杂的计算机系统中单纯使用软件的方法难以解决所有的问题,一种可以尝试的方向就是利用硬件芯片和可信计算的支持,在云的环境中建立可信计算基(TCB)保护用户、基础设施提供商、服务提供商的秘密,进行完整性度量以及执行云计算参与各方的身份证明和软件可信性证明,也即构造基于可信基的可信云计算(Trusted Cloud Computing).

EMC 中国实验室与复旦大学、华中科技大学、清华大学、武汉大学协作,开展道里可信虚拟基础设施研究项目. 该研究项目致力于在云计算的多租户计算环境中,实现租户隔离,保护平台提供者不受恶意租户的攻击. 结合可信计算和虚拟化技术来加强计算平台的安全,使得云服务商能够在公共云计算平台中提供虚拟私有云计算服务(Virtual Private Cloud, VPC). 虚拟私有云的实现需要对云服务提供者的内存储器和 CPU 寄存器作一种非加密方式的保护,使得租客的代码和数据在云服务提供者的内存和 CPU 寄存器中以明文形式被处理时仍然得到私密性及完整性的保护,避免被其它租客或攻击者窃取. 项目提供的 VPC 计算服务为云用户提供应用程序级别的安全隔离,并保证用户代码和数据的私密性和完整性,是从真正意义上降低了云计算的安全风险[51].

文献[52]讨论了虚拟机管理平台 Xen 中,利用 TCB 的安全增强措施,描述了这种方法如何被用于实现"可信虚拟化"及提高虚拟 TPM 实现的安全性. 目前 Xen 的 TCB 除了虚拟机监控器 VMM 外,还包含一个完整的 OS(Dom0)和一个用户空间工具集合. 这使得 TCB 异常笨重. 并且用户空间工具集的存在也使得硬件管理员能配置任意特权代码到 TCB 中,带来不安全因素. 本文把新的 VM 创建功能转移到一个小的运行于 Dom0 之外的可信 VM,这样做有两个主要目标:基本目标是减小和界定基于 Xen 的系统的 TCB,尤其是将 Dom0 用户空间从 TCB 中移除,从而提高安全性. 另一个目标是如果假设 TCB 安全的话,那么新创建的 VM 保持了与物理机器一样的安全和完整属性.

文献[31]介绍的隐私管理器中也利用了 TPM 来管理隐私保护过程中所需的密钥.

文献[53]将可信计算技术引入到 IaaS 类型的云计算体系中,以开源的 IaaS 平台 Eucalyptus 为例,引入一个可信协调器(TC)(由一个外部可信实体(ETE)来维护),将不可信的云管理器(CM)与若干可信的节点结合起来形成其总体架构. 文中还详细介绍了在这样的架构下的可信节点的生成和管理,虚拟机的管理和迁移等问题,并提出了在完成这些工作期间需要的信息交换协议. 这种 TCCP 架构保证了客户 VM 的安全执行,允许用户对 IaaS 服务提供商进行验证以及在启动 VM 之前确定服务是否安全. 他们的下一步工作是实现一个完整的原型系统并进行性能评估.

7　总结和展望

然而云计算安全相关的研究仍处于起步阶段,许多问题仍待探索. 本文总结了以前的云计算相关的安全研究成果,认为一个重要的研究方向就是将可信计算技术应用在云中. 然而,根据 TCG 的可信平台规范 v1.2,当前 TPM 芯片以串行指令方式进行处理,其计算速度,密钥存储空间以及计算性能是有限的. 并且,移动计算和无线网络的发展也要求应用云服务所带来的便利. 所以,未来研究的方向可以考虑如下几个方面:

- 在移动计算中建立可信云计算的基础结构和应用.
- 在云计算中通过使用TPM进行远程认证的相关协议.
- 在云中利用用户的数字 ID 和可信计算技术,进行用于认证的密钥管理技术和建立层次型密钥管理结构,同时与基于身份的加密技术结合,从而提高云端数据的安全性,保密性和完整性.
- 用于云计算环境中的 TPM 使用的密码算法的发展,例如椭圆曲线密码.
- 基于 TPM 和可信计算基的公共 PKG 基础设施建设.
- 加密数据上的操作和运算方法研究(例如使用同型哈

希函数,代理重加密和安全多方计算的概念).

References:

[1] Brodkin. Gartner: seven cloud-computing security risks[DB/OL]. http://www.networkworld.com/news/2008/070208-cloud.html, 2008-07-02.

[2] Cloud Security. Alliance security guidance for critical areas of focus in cloud computing V2.1[EB/OL]. http://www.cloudsecurityalliance.org/guidance/csaguide.v2.1.pdf,2009-12-01.

[3] Meiko Jensen, Jorg Schwenk, Nils Gruschka, et al. On technical security issues in cloud computing[C]. Cloud, 2009 IEEE International Conference on Cloud Computing, 2009:109-116.

[4] Francis Hsu, Chen Hao. Secure file system services for Web 2.0 applications[C]. In CSW'09: Proceedings of the 2009 ACM Workshop on Cloud Computing Security, New York, NY, USA, 2009:11-18.

[5] Xu Jin-song, Huang Ru-cheng, Huang Wan-ming, et al. Secure document service for cloud computing[C]. In CloudCom'09: Proceedings of the 1st International Conference on Cloud Computing, Beijing, China, 2009:541-546.

[6] Cao Tian-jie, Zhang Yong-ping, Wang Chu-jiao. Security protocol [M]. Beijing: Beijing University of Posts and Telecommunications Press, 2009.

[7] Aaram Yun, Shi Chun-hui, Yongdae Kim. On protecting integrity and confidentiality of cryptographic file system for outsourced storage[C]. In CCSW'09: Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA, 2009:67-75.

[8] Carter J L, Wegman M N. Universal classes of hash functions (extended abstract)[A]. STOC '77: Proceedings of the Ninth Annual ACM Symposium on Theory of Computing, New York, NY, USA [C]. ACM Press, 1977:106-112.

[9] Wegman M N, Carter L. New classes and applications of hash functions[C]. In FOCS, IEEE, 1979:175-182.

[10] Wang Cong, Wang Qian, Ren Kui, et al. Ensuring data storage security in cloud computing[C]. In IWQoS'09: Proceedings of 17th International Workshop on Quality of Service, Charleston, SC, USA, 2009:1-9.

[11] Adi Shamir. Identity-based cryptosystems and signature schemes [C]. In Proceedings of CRYPTO'84 on Advances in Cryptology, Santa Barbara, California, USA, 1985:47-53.

[12] Gentry C, Silverberg A. Hierarchical ID-based cryptography[C]. In Proceedings of ASIACRYPT 2002, LNCS, Springer, Heidelberg,2002,2501:548-566.

[13] Horwitz J, Lynn B. Toward hierarchical identity-based encryption [C]. In Proceedings of EUROCRYPT 2002, LNCS, Springer, Heidelberg,2002,2332:466-481.

[14] Yan Liang, Rong Chun-ming, Zhao Gan-sen. Strengthen cloud computing security with federal identity management using hierarchical identity-based cryptography[C]. In CloudCom'09: Proceedings of the 1st International Conference on Cloud Computing, Beijing, China, 2009:167-177.

[15] Li H, Dai Y, Tian L, et al. Identity-based authentication for cloud computing[C]. In CloudCom'09: Proceedings of the 1st International Conference on Cloud Computing, Beijing, China, 2009: 157-166.

[16] Hu L, Ying S, Jia X, et al. Towards an approach of semantic access control for cloud computing[C]. In CloudCom'09: Proceedings of the 1st International Conference on Cloud Computing, Beijing, China, 2009:145-156.

[17] Yu S, Wang C, Ren K, et al. Achieving secure, scalable, and fine-grained data access control in cloud computing[C]. In Proceedings of IEEE INFOCOM, 2010:534-542.

[18] Goyal V, Pandey O, Sahai A, et al. Attribute-based encryption for fine-grained access control of encrypted data[C]. In Proc. of CCS'06, 2006.

[19] Blaze M, Bleumer G, Strauss M. Divertible protocols and atomic proxy cryptography[C]. In Proc. of EUROCRYPT '98, 1998.

[20] Wang W, Li Z, Owens R, et al. Secure and efficient access to outsourced data[C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA, 2009:55-66.

[21] Atallah M J, Blanton M, Fazio N, et al. Dynamic and efficient key management for access hierarchies[J]. ACM Trans. Inf. Syst. Secur.,2009,12(3):1-43.

[22] Damiani E, di Vimercati S D C, Foresti S, et al. Key management for multi-user encrypted databases[C]. In Proceedings of the ACM Workshop on Storage Security and Survivability, 2005:74-83.

[23] di Vimercati S D C, Foresti S, Jajodia S, et al. Over-encryption: management of access control evolution on outsourced data[C]. In Proceedings of the International Conference on Very large Data Bases, 2007:123-134.

[24] Kallahalla M, Riedel E, Swaminathan R, et al. Plutus: scalable secure file sharing on untrusted storage[C]. In Proceedings of the USENIX Conference on File and Storage Technologies, 2003:29-42.

[25] Wei Jin-peng, Zhang Xiao-lan, Glenn Ammons. Managing security of virtual machine images in a cloud environment[C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, 2009, Chicago, Illinois, USA, 2009:91-96.

[26] Matthews J, Garfinkel T, Hoff C, et al. Virtual machine contracts for datacenter and cloud computing environments[C]. In Workshop on Automated Control for Datacenters and Clouds (ACDC), ACM, 2009:25-30.

[27] Christodorescu M, Sailer R, Schales D L, et al. Cloud security is not (just) virtualization security[C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA, 2009:97-102.

[28] Raj H, Nathuji R, Singh A, et al. Resource management for isolation enhanced cloud services[C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA, 2009:77-84.

[29] Dominik Herrmann, Rolf Wendolsky, Hannes Federrath. Website fingerprinting: attacking popular privacy enhancing technologies with the multinomial naive- bayes classifier[C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA, 2009:31-41.

[30] Mowbray M, Pearson S. A client-based privacy manager for cloud

computing[C]. In Proceedings of the 4th International ICST Conference on Communication System Software and Middleware, Dublin, Ireland, 2009:1-8.

[31] Siani Pearson, Shen Yun, Miranda Mowbray. A privacy manager for cloud computing[C]. In CloudCom'09: Proceedings of the 1st International Conference on Cloud Computing, Beijing, China, 2009:90-106.

[32] Liu Q, Wang G, Wu J. An efficient privacy preserving keyword search scheme in cloud computing[C]. In Proceedings of IEEE 2009 International Conference on Computational Science and Engineering/TrustCom, 2009:715-720.

[33] Li J, Wang Q, Wang C, et al. Fuzzy keyword search over encrypted data in cloud computing[C]. In Proc. of IEEE INFOCOM'10 Mini-Conference, San Diego, CA, USA, 2010.

[34] Mariana Raykova, Binh Vo, Steven M Bellovin. Secure anonymous database search[C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA, 2009:115-126.

[35] Burton H Bloom Space/time trade-offs in hash coding with allowable errors[J]. Communications of the ACM, 1970,13(7):422-426.

[36] Matt Blaze, Gerrit Bleumer, Martin Strauss. Divertible protocols and atomic proxy cryptography[C]. In Proceedings of EUROCRYPT'98, 1998.

[37] Josh Benaloh, Melissa Chase, Eric Horvitz, et al. Patient controlled encryption: ensuring privacy of electronic medical records [C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA, 2009:103-114.

[38] Pearson S. Taking account of privacy when designing cloud computing services[C]. In Proceedings of ICSE-Cloud'09, Vancouver, Canada IEEE, 2009.

[39] Ateniese G, Burns R, Curtmola R, et al. Provable data possession at untrusted stores[A]. In Proceedings of CCS'07[C], ACM Press, New York, 2007:598-609.

[40] Ateniese G, Di Pietro R, Mancini L V, et al. Scalable and efficient provable data possession[C]. In Proceedings of Secure Comm'08, 2008.

[41] Wang Q, Wang C, Li J, et al. Enabling public verifiability and data dynamics for storage security in cloud computing[C]. In Proceedings of ESORICS'09, Saint Malo, France, 2009.

[42] Shacham H, Waters B. Compact proofs of retrievability[C]. In Proceedings of ASIACRYPT 2008, LNCS, Springer, Heidelberg, 2008,5350:90-107.

[43] Juels A, Kaliski Jr B S. Pors: proofs of retrievability for large files [A]. In Proceedings of CCS'07[C], ACM Press, New York, 2007:584-597.

[44] Wang C, Wang Q, Ren K, et al. Privacy-preserving public auditing for data storage security in cloud computing[C]. In Proceedings of IEEE INFOCOM'10, 2010.

[45] Kevin D Bowers, Ari Juels, Alina Oprea. Proofs of retrievability: theory and implementation[C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA, 2009:43-53.

[46] Du J, Wei W, Gu X, et al. Runtest: assuring integrity of dataflow processing in cloud computing infrastructures[C]. In ACM Symposium on Information, Computer and Communications Security (ASIACCS), 2010.

[47] Lu Rong-xing, Lin Xiao-dong, Liang Xiao-hui, et al. Secure provenance: the essential of bread and butter of data forensics in cloud computing[C]. In Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security, Beijing, China, 2010:282-292.

[48] Zhang Xin-wen, Joshua Schiffman, Simon Gibbs. Securing elastic applications on mobile devices for cloud computing[C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA, 2009:127-134.

[49] Huang Wei-li, Yang Jian. New network security based on cloud computing[C]. In Proceedings of the Second International Workshop on Education Technology and Computer Science, 2010.

[50] Robert Biddle P C, van Oorschot, Andrew S Patrick. Browser interfaces and extended validation SSL certificates: an empirical study[C]. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA, 2009:19-30.

[51] Mao Wen-bo. Also talking about the cloud[EB/OL]. http://blog.csdn.net/wenbomao/archive/2009/03/03/3952761.aspx, 2009-03-03.

[52] Murray D G, Milos G, Hand S. Improving xen security through disaggregation[C]. In Proceedings of VEE'08, New York, NY, USA, 2008:151-160.

[53] Santos N, Gummandi K P, Rodrigues R. Towards trusted cloud computing[C]. In Workshop on Hot Topics in Cloud Computing, San Diego, CA, 2009.

附中文参考文献:

[6] 曹天杰, 张永平, 汪楚娇. 安全协议[M]. 北京: 北京邮电大学出版社, 2009.

[51] 毛文波. 我亦云云-也谈云计算[EB/OL]. http://blog.csdn.net/wenbomao/archive/2009/03/03/3952761.aspx, 2009-03-03.