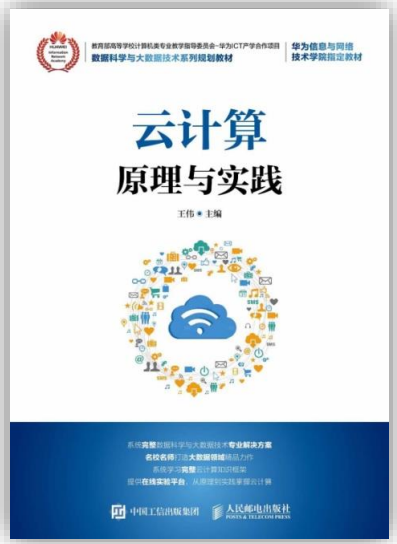
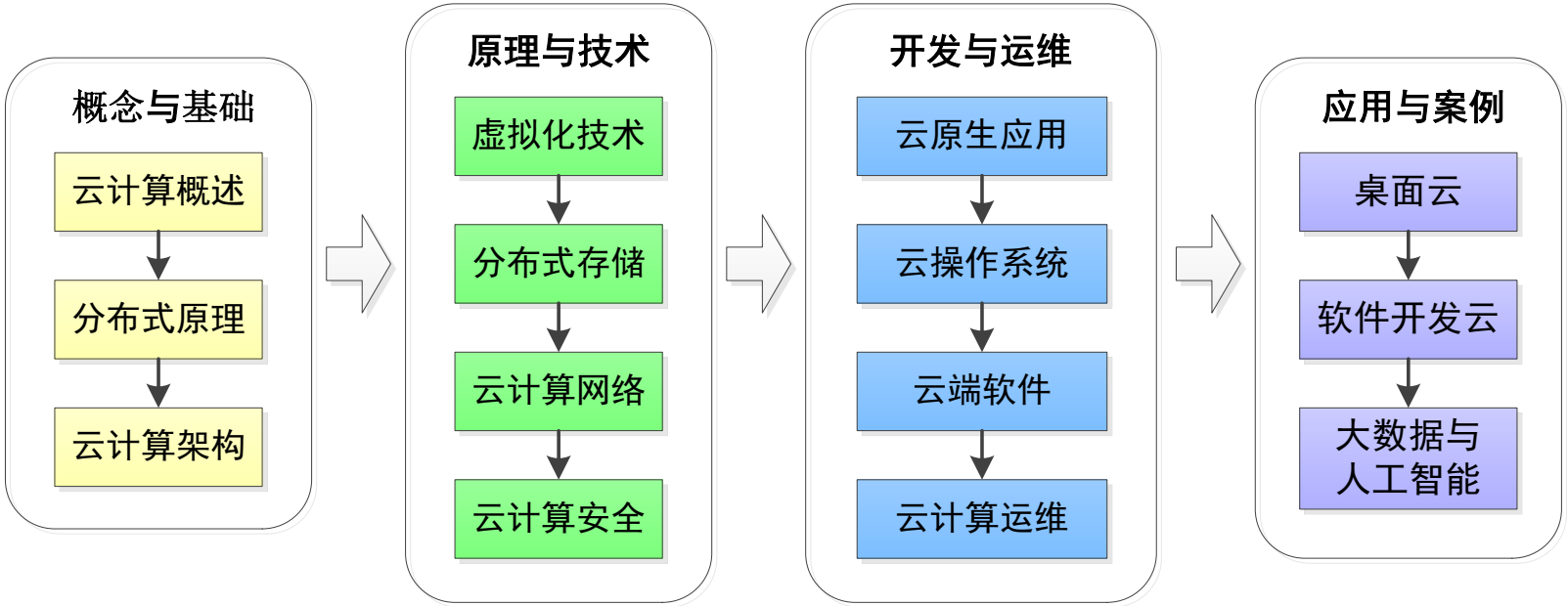

云计算原理与实践
Principles and Practice of Cloud Computing

云计算原理与实践

Principles and Practice of Cloud Computing



《云计算原理与实践》课程总览



Outline

- 7.1 云安全概述
- 7.2 虚拟机安全
- 7.3 云存储安全
- 7.4 云数据安全
- 7.5 实践：全同态加密算法

7.1 云安全概述

1. 云计算安全挑战
2. 云计算安全现状
3. 云计算安全技术框架
4. 云计算安全关键技术

7.1.1 云计算安全挑战

在云计算安全上一直有这样一种分歧。一方认为，采用云计算能够增强安全性，通过部署集中的云计算中心，可以组织安全专家及专业化安全服务队伍实现整个系统的安全管理，避免由个人维护安全及不专业导致安全漏洞频出而被黑客利用的情况。另一方则持反对意见，认为集中管理的云计算中心将成为黑客攻击的重点目标。

7.1.2 云计算安全现状

1．各国政府对云计算安全的关注

2．云计算安全标准组织及其进展

各国政府对云计算安全的关注

2010年3月，参加欧洲议会讨论的欧洲各国网络法律专家和领导人呼吁制定一个关于数据保护的全球协议，以解决云计算的数据安全弱点。欧洲网络和信息安全局（ENISA）表示，将推动管理部门要求云计算提供商必须通知客户有关安全攻击状况。

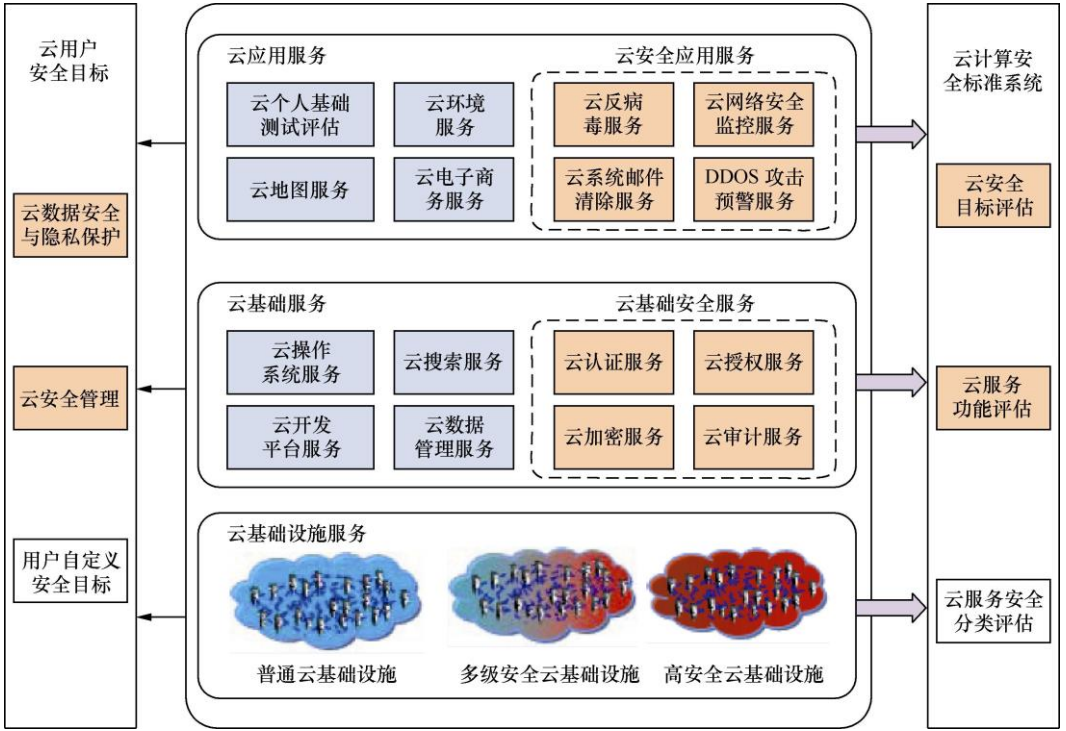
2010年11月，美国CIO委员会发布关于政府机构采用云计算的政府文件，阐述了云计算带来的挑战以及针对云计算的安全防护，要求政府机构评估云计算相关的安全风险并与自己的安全需求进行比对分析。

我国从2010年开始，加强云计算信息安全研究，解决共性技术问题，保证云计算产业健康、可持续地发展。

云计算安全标准组织及其进展

- 国际电信联盟ITU-TSG17研究组
- 结构化信息标准促进组织与分布式管理任务组（Distributed Management Task Force，DMTF）
- ITU-TSG17研究组
- 云安全联盟（Cloud Security Alliance，CSA）

7.1.3 云计算安全技术框架



7.1.4 云计算安全关键技术

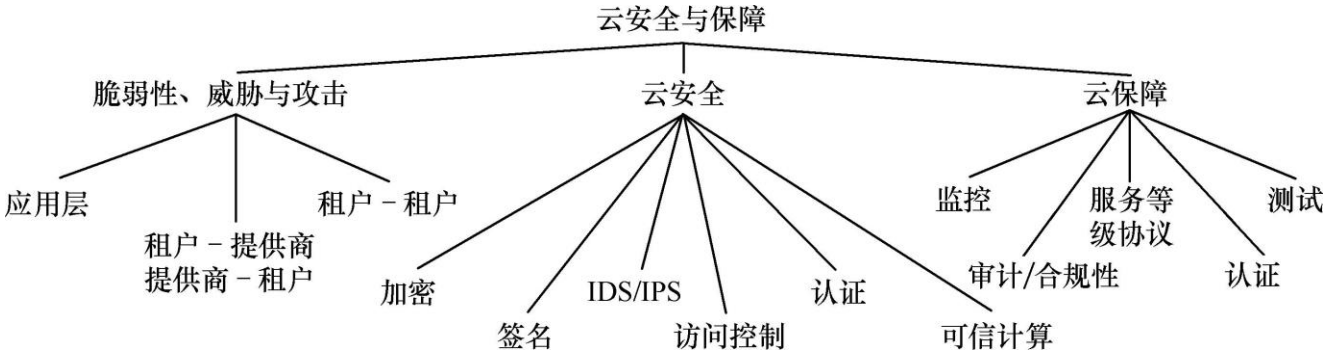


图7.2 云安全与保障的技术体系

云计算安全需求的重点

- 可信访问控制
- 密文检索与处理
- 数据存在与可使用性证明
- 数据隐私保护
- 虚拟安全技术
- 云资源访问控制
- 可信云计算

7.2 虚拟机安全

1. 虚拟化软件栈安全威胁
2. 虚拟化软件栈安全防御
3. 虚拟化安全总结

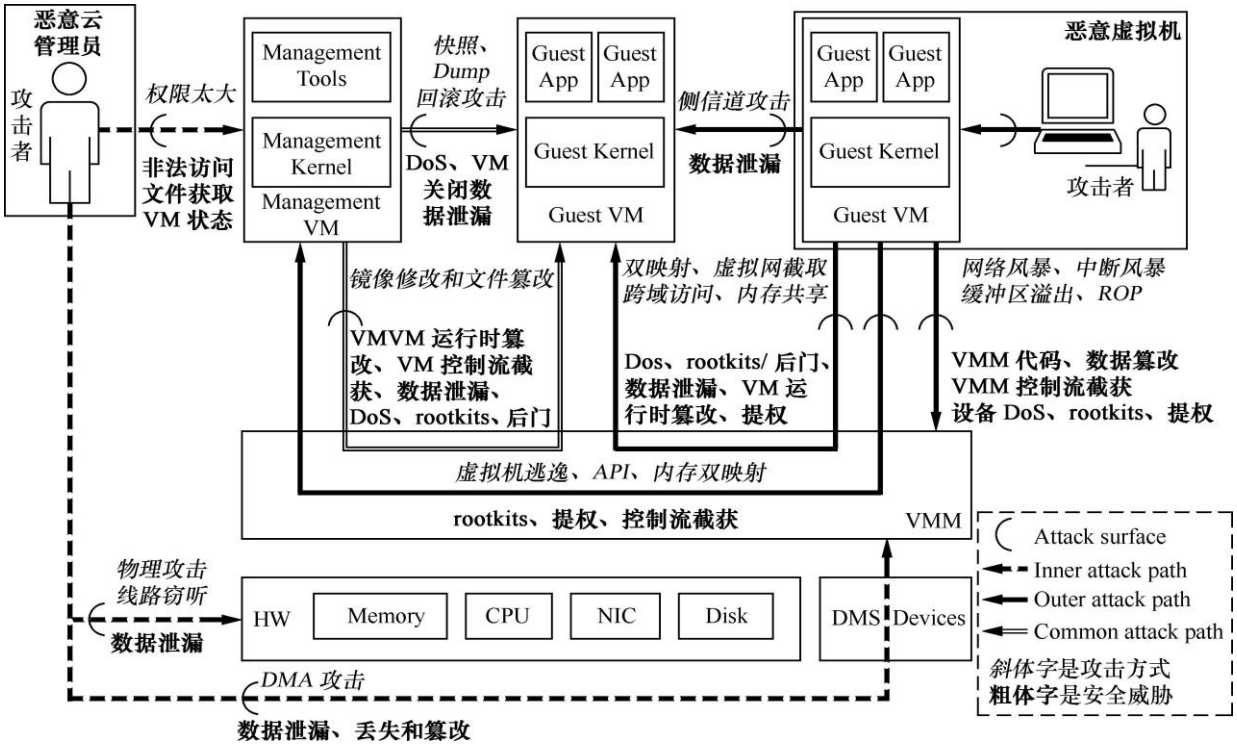
7.2.1 虚拟化软件栈安全威胁

1．攻击来源

2．安全威胁

3．攻击方式

1 · 攻击来源


恶意云
管理员
攻击者
权限太大
非法访问
文件获取
VM 状态
Management
Tools
Management
Kernel
Management
VM
快照、
Dump
回滚攻击
DoS、VM
关闭数
据泄漏
Guest
App
Guest
App
Guest Kernel
Guest VM
侧信道攻击
数据泄漏
Guest
App
Guest
App
Guest Kernel
Guest VM
恶意虚拟机
攻击者
镜像修改和文件篡改
VMVM 运行时篡
改、VM 控制流截
获、数据泄漏、
DoS、rootkits、后门
双映射、虚拟网截取
跨域访问、内存共享
Dos、rootkits/ 后门、
数据泄漏、VM 运
行时篡改、提权
网络风暴、中断风暴
缓冲区溢出、ROP
VMM 代码、数据篡改
VMM 控制流截获
设备 DoS、rootkits、提权
虚拟机逃逸、API、内存双映射
rootkits、提权、控制流截获
VMM
物理攻击
线路窃听
数据泄漏
HW
Memory
CPU
NIC
Disk
DMS
Devices
DMA 攻击
数据泄漏、丢失和篡改
Attack surface
Inner attack path
Outer attack path
Common attack path
斜体字是攻击方式
粗体字是安全威胁

2．安全威胁

- 威胁1：数据泄露和丢失
- 威胁2：控制流截获及后门、rootkits
- 威胁3：拒绝服务（Denial of Service，DoS）
- 威胁4：虚拟机镜像威胁
- 威胁5：运行时代码、数据篡改
- 威胁6：权限提升
- 威胁7：不可信的云内部人员

3．攻击方式

- DMA攻击
- 多重映射和虚拟机跨域访问
- 跨虚拟机的Cache攻击
- 快照、内存转存威胁
- 物理攻击和线路窃听

DMA攻击

DMA的初衷是允许外围设备绕过MMU，直接对物理内存进行读写操作，从而提高I/O效率。在Intel VTd提出之前，具有DMA功能的外设可以对物理内存进行任意访问，VTd的提出使这一问题得到了缓解。DMA攻击主要分为3步：首先，对外设进行改造，嵌入恶意代码；然后，将外设部署到目标主机中；最后，利用恶意代码发送DMA请求，实现恶意攻击。DMA攻击的难点是定位需要访问的数据结构或代码的地址，如此才能精确地实现有目的的攻击。否则，只能利用DMA进行粗粒度的数据窃取。在虚拟化场景下，内部攻击者可以通过DMA设备，对物理内存中的代码、数据进行篡改或窃取，从而实现代码注入、控制流劫持和数据泄露等。当前的主要解决方案是结合IO MMU对DMA的读写操作进行限制。

多重映射和虚拟机跨域访问

跨域访问是指客户虚拟机不仅能够访问自身的地址空间，同时还能够访问到其他虚拟机或Hypervisor地址空间中的数据。在IaaS模型中，每个 虚拟机都有独立的EPT（Extend Page Table）或SPT（Shadow Page
Table），并且Hypervisor拥有单独的地址空间。然而，攻击者利用一 些软件漏洞、DMA攻击、VLAN跳跃攻击和Cache变更等实现虚拟机跨 域访问。例如，攻击者利用Hypervisor漏洞或者已控制的Hypervisor对 客户虚拟机的页表进行修改，使其映射到另一客户虚拟机的地址空间中，从而实现跨域访问。跨域访问能够窃取或篡改其他用户的数据或建立隐蔽信道。防止这类攻击的主要方式是对不同虚拟机之间进行隔离，并且剥夺Hypervisor更新EPT页表的能力。

跨虚拟机的Cache攻击

随着计算模式从独占计算硬件到云计算模式的迁移，基于共享Cache的侧信道攻击变得越发严重。基于Cache的侧信道攻击和隐蔽信道攻击使攻击者能够在数秒或数分钟内从当前流行的加密方法（RSA、AES和DES）中获取到受害者的密钥信息。基于Cache的侧信道攻击不需要获取Hypervisor等特权和利用其漏洞，而只需通过对时间损耗、电源损耗及电磁辐射等特性的监测、统计即可获取到其他客户虚拟机的数据。侧信道攻击可以分为3种方式：基于时间驱动、基于轨迹驱动和基于访问驱动。基于时间驱动的攻击是攻击者重复地检测被攻击者的加密操作所使用的时间，然后通过差分分析等技术推断出密钥等信息。基于轨迹驱动的攻击通过持续地对设备的电能损耗、电磁发射等情况进行监控，获取到其敏感信息，但是这类侧信道攻击需要攻击者能够物理接近攻击目标。基于访问驱动的攻击是攻击者在执行加密操作的系统中运行一个应用，这个应用通过监控共享Cache的使用情况来获取密钥信息。基于访问驱动攻击的优势是不需要攻击者得到受害者精确的时间信息。

快照、内存转存威胁

虚拟机快照（snapshot）是Hypervisor提供管理者的API，用于容错和虚拟机维护。云提供商的内部管理员可以利用管理工具对运行中的虚拟机进行快照，为内部攻击者提供了便利之门。这样可以在用户不知情的情况下，就可将虚拟机回滚（rollback）到特定阶段，从而绕过一些安全机制的更新。内部攻击者甚至可以利用内存转存工具对用户的内存进行转储，然后进行线下分析、窃取用户数据。通常这类攻击的防护是利用密码学机制防护，或者禁用管理员的快照和转存操作。

物理攻击和线路窃听

物理攻击是指攻击者能够物理接近攻击目标所在的物理服务器。虽然数据中心有专门的安全防护措施（例如录像监控和审计机制），但是数据中心的机房每天都有维修人员、清洁人员和管理人员出入，给安全带来了潜在的隐患。冷启动攻击就是很好的例子。通道或线路窃听可认为是另一种形式的物理攻击，攻击者通过一些特殊的方式监听受害者的通道和线路，包括外部网络、虚拟机之间的虚拟网络和内部总线等，从中窃取来自或流向虚拟机和Hypervisor的数据。

7.2.2 虚拟化软件栈安全防御

虚拟化软件栈安全可分为虚拟机自身的（GOS、Apps）的安全和Hypervisor（虚拟化层）的安全两个层次。从可信基的角度分类，业界的安全方案可分为基于Hypervisor的保护、Hypervisor自身安全防护及虚拟机在不可信Hypervisor环境中的安全防护。其他方案还包括在Hypervisor层之下引入新的软硬件安全模块，从隔离机制、加密机制和权限访问控制这些不同角度对虚拟机及内部软件进行保护，以及对侧信道攻击的防护。

7.2.3 虚拟化安全总结

近年来虚拟化安全虽取得众多成果，但针对现存的安全问题仍捉襟见肘，针对众多潜在的安全威胁和漏洞更是力不从心。由此可见，虚拟化安全还有很大的研究和提升空间，迫切需要研究出一套高效、可行且易实施的虚拟化安全防护方案。

7.3 云存储安全

1. 云存储的安全需求
2. 安全云存储系统概述
3. 安全云存储系统的一般架构
4. 安全云存储系统的关键技术

7.3.1 云存储的安全需求

1．数据的安全性
2．密钥管理分发机制
3．其他功能需求

7.3.2 安全云存储系统概述

用户对云存储的不信任引发了云存储系统中的安全问题。近年来，随着云存储的推广与普及，虽然有越来越多的人开始使用云存储存放自己的资料，但云存储系统中的安全问题却并没有得到缓解。为了解决云存储系统中的安全问题，国内外的研究者做了大量研究，逐渐在云存储系统的研究中形成一个新的方向——安全云存储系统。

安全云存储系统设计的一般原则

安全云存储系统是云存储系统的一个子集，它指的是包含安全特性的云存储系统，安全云存储系统的设计者通常会提出一些安全方面的假设，然后根据这些假设建立系统的威胁模型与信任体系，最终设计并实现系统或原型系统。

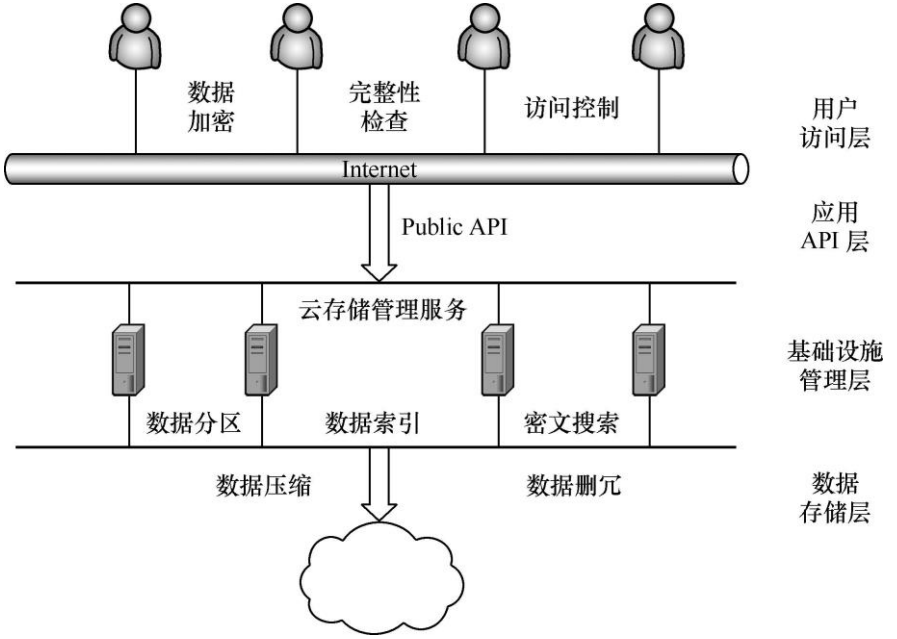
安全云存储系统的现状

从存储系统的技术支撑与发展来看，文件系统是构建云存储系统的重要部分之一。随着网络存储系统的发展，加密文件系统的理念也逐渐网络化、系统化，最终演变成安全网络存储系统。一般的安全网络存储系统至少包括客户端与服务器两部分，客户端由系统的使用者进行操作，为用户数据提供数据加解密、完整性校验以及访问权限控制等功能；服务器作为数据及元数据的存储介质，对数据没有任何的访问或使用权限。

7.3.3 安全云存储系统的一般架构

云存储按照其体系结构可分为存储层、基础管理层、应用接口层和访问层。在具体的安全云存储系统中，由于应用场景和研究目标的不同，其系统架构也各不相同。图7.4总结归纳了现有安全云 存储系统的通用架构，具体的安全云存储系统只 需根据自身的特点实现部分或全部的功能。

安全云存储系统的通用架构



7.3.4 安全云存储系统的关键技术

1．安全、高效的密钥生成管理分发机制
2．基于属性的加密方式
3．基于密文的搜索方式
4．基于密文的重复数据删除技术
5．基于密文的数据持有性证明
6．数据的可信删除

1．安全、高效的密钥生成管理分发机制

在目前的安全云存储系统中，数据加密存储是解决机密性问题的主流方法。数据加密时必须用到密钥，在不同系统中，根据密钥的生成粒度不同，需要管理的密钥数量级也不一样。若加密粒度太大，虽然用户可以很方便地管理，却不利于密钥的更新和分发；若加密粒度太小，虽然用户可以进行细粒度的访问权限控制，但密钥管理的开销也会变得非常大。

2．基于属性的加密方式

在公私钥加密体系中有一种特殊的加密方式：基于属性的加密方式（Attribute-based Encryption）。基于属性的加密方式以属性作为公钥对用户数据进行加密，用户的私钥也和属性相关，只有当用户私钥具备解密数据的基本属性时，用户才能够解密出数据明文。

3．基于密文的搜索方式

目前可搜索加密机制的研究可分为基于对称加密（Symmetry Key Cryptography Based）的SK机制和 基于公钥加密（Public Key Cryptography Based）的
SE机制两类。

4．基于密文的重复数据删除技术

在一般的云存储系统中，为了节省存储空间，系统或多或少会采用一些重复数据删除（Data Deduplication）技术来删除系统中的大量重复数据。但是在安全云存储系统中，与数据搜索问题一样，相同内容的明文会被加密成不同的密文，因此也无法根据数据内容对其进行重复数据删除操作。

5．基于密文的数据持有性证明

目前的数据持有性证明主要有可证明数据持有（Provable Data Possession，PDP）和数据证明与恢复（Proof OF Retrievability，POR）两种方案。

6．数据的可信删除

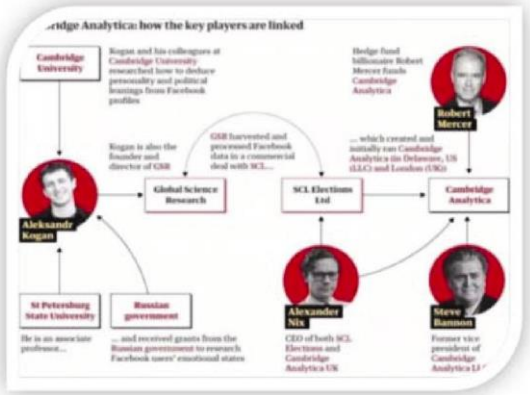
云存储的可靠性机制在提高数据可靠性的同时也为数据的删除带来了安全隐患：数据存储在云存储中，当用户向云存储下达删除指令时，云存储可能会恶意地保留此文件，或者由于技术原因并未删除所有副本。一旦云存储通过某种非法途径获得数据密钥，数据也就面临着被泄露的风险

7.4 云数据安全

1. 云数据面临的安全威胁
2. 云数据安全研究内容
3. 云数据安全研究进展

7.4.1 云数据面临的安全威胁

推送广告
影响大选



泄露用户数据

扎克伯格公开道歉

Cambridge Analytica

图7.5 Facebook大规模数据泄露事件

7.4.2 云数据安全研究内容

云数据服务	安全威胁	安全需求	研究内容	
云数据存储	数据破坏或丢失	数据完整性	云数据安全验证	支持数据动态操作的验证 公开可审 计验证 数据可恢复证明
云数据共享	非法访问	访问可控性	云数据安全共享	细粒度访问控制 访问权限动态更新 用户动态添加或撤销
云数据查询	数据泄露	数据机密性	云数据安全查询	支持丰富的查询功能 支持数据动态 变化 支持查询结果排序
云数据计算	数据泄露	数据机密性	云数据安全查询	支持密文计算的同态加密 特定类型 安全外包计算 外包计算结果验证

7.4.3 云数据安全研究进展

1．云数据安全验证研究进展
2．云数据安全共享研究进展
3．云数据安全查询研究
4．云数据安全计算研究

1．云数据安全验证研究进展

在云存储环境下，由于用户带宽和存储资源的限制，用户不可能将数据全部取回进行完整性验证。因此，传统的数据完整性验证方案无法直接应用于云数据验证场景。针对此问题，研究者们提出了支持远程验证技术，即在不下载验证数据的前提下，仅通过简单的挑战—应答方式来完成数据的验证。典型的方案是数据持有性证明（Provable Data Possession，PDP），该方案采用同态验证标签，具有聚合特性，能够将多块数据验证的证据聚合为一个验证响应，降低验证响应的带宽消耗。数据持有性证明方案还采用了随机抽样的概率性验证方法，有效降低了验证通信和计算开销。解决这方面问题的技术包括：支持动态数据操作、支持公开可验证及数据可恢复性证明等。

2．云数据安全共享研究进展

在不可信云环境下，数据拥有者通常会将数据加密后上传到云服务器中，但这给数据共享带来一定的困难。一是大规模用户的数据共享需要大量密钥，生成、分发和保管这些密钥比较困难；二是如果制定灵活可控的访问策略，实施细粒度的访问控制，会成倍地增加密钥数量；三是当用户访问权限更新或撤销时，需要重新生成新的密钥，势必引入巨大的计算量。另一个重要问题是传统的访问控制方法依赖于一个可信的服务器，而该假设条件在不可信云计算环境下是不成立的。解决这方面问题的技术包括：基于属性的加密技术、访问策略表达技术、访问权限撤销技术及访问控制效率增强技术等。

3．云数据安全查询研究

出于数据机密性的考虑，用户会将数据以密文的形式外包存储在云服务器中。但当用户需要提取包含某些关键字的数据时，会遇到如何在云端服务器进行密文搜索的难题。一种简单的方法是将所有密文数据下载到本地进行解密，然后再进行关键字查询，但这种方法会浪费巨大的网络带宽，给用户带来大量不必要的存储和计算开销。另一种方法是将密钥和要查询的关键字发给云服务器，由云服务器解密数据后进行查询，但这种方法会泄露用户数据，不能满足数据机密性要求。为此，支持密文搜索的查询加密（Searchable Encryption, SE）技术应运而生，其基本思想是通过构造安全索引、利用查询陷门来高效地支持密文搜索。

4．云数据安全计算研究

在解决大规模最优化、大数据分析、生物特征匹配等问题时，会涉及大量的数据计算。对于资源有限的用户来说，承担如此巨大的计算比较困难。一种有效的解决方案是借助云端的强大计算能力为用户提供计算服务，但这会将用户的敏感数据暴露给云服务器。解决的方法是可以通过加密数据，让云服务器在密文数据上进行计算。另一个重要问题是在不可信的云环境下，云服务器是否能够正确可靠地为用户进行所需的计算，并返回正确的结果，用户不得而知。这可以通过研究可验证外包技术来解决。解决这方面问题的技术包括：同态加密、特定计算安全外包及可验证外包计算等。

7.5 实践：全同态加密算法

1. HElib库的调试与分析
2. FHE-CODE的调试与分析
3. 全同态加密方案对比与分析

7.5.1 HElib库的调试与分析

HElib库是基于C++语言的同态加密算法软件库。当前，能够实现的是BGV同态加密体制（由Brakerski、Gentry和Vaikuntanathan三位联合提出的全同态加密算法），以及许多为了提高算法运行速度和着重聚焦使用SV（Smart-Vercauteren）的密文封装技术和GHS（Gentry-Halevi-Smart）的优化算法.

HElib库的调试过程

- 安装NTL算法库。由于NTL算法库是基于GMP大数库编写的，因此首先确定系统已正确安装GMP大数库。其次，通过Linux的gunzip、tar、make、make install等命令执行NTL安装过程。
- 调试阶段。进入HElib-src文件中，对文件执行make操作，此时编译产生了许多的.o文件，与此同时还可以在src文件夹中发现程序在运行的过程中自动生成了fhe.a的静态链接库。
- 从Makefile文件中可以得知要通过make check产生可执行文件，与此同时可以查看src文件夹中的变化，并且实现程序的运行以及相关数据的检测。

7.5.2 FHE-CODE的调试与分析

FHE-CODE是Gentry的全同态加密方案的一种变体，适用于64位的Intel处理器。该算法库所依据的算法包括用于密钥生成算法、加密算法、解密算法、密文刷新算法。

FHE-CODE的调试过程

首先，对程序代码进行编译工作。由于FHE-CODE代码库 的编写与HElib库的编写基于相同的链接库，所以可以基 于之前搭建的环境直接进行调试。使用make命令对FHE-CODE文件进行编译即可。编译后，在当前目录下会产生.o格式的目标文件，这说明编译成功

7.5.3 全同态加密方案对比与分析

HElib的分析

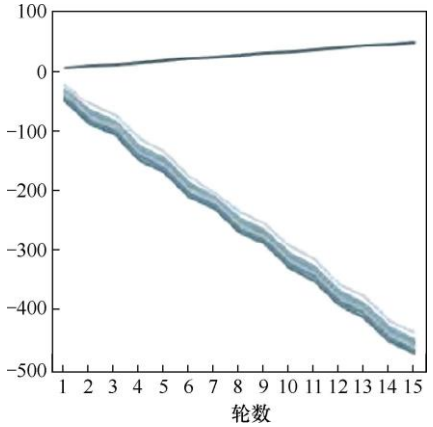


图7.14 不同轮数第一轮噪声与模数比 变化关系

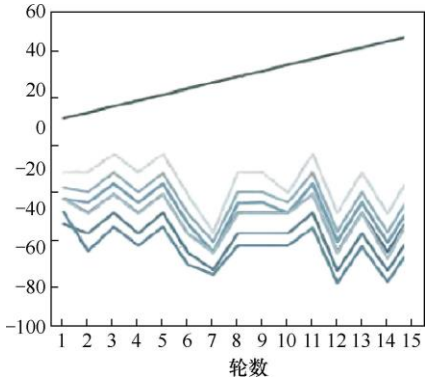


图7.15 不同轮数最后一轮噪声与模数比对数、级数的对数、级数的变化关系

7.5.3 全同态加密方案对比与分析

FHE-code的分析

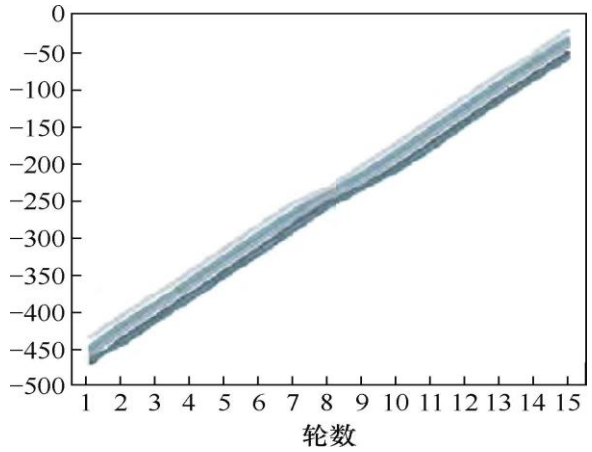


图7.16 轮数为15时每一轮噪声与模数比对数的变化关系

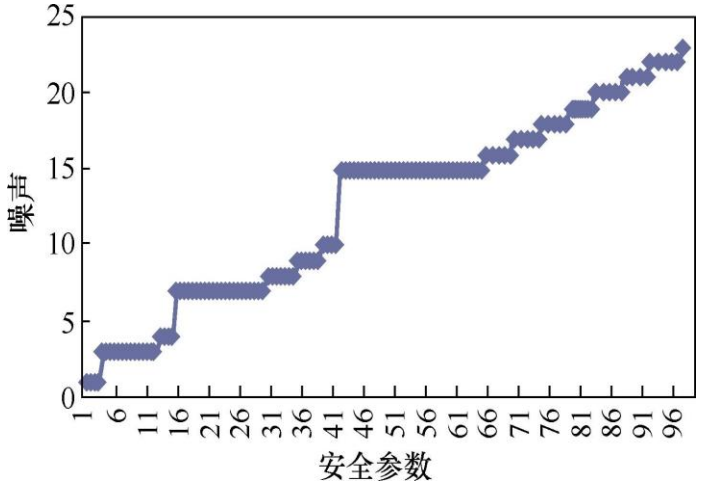


图7.17 噪声与安全参数的关系

7.5.3 全同态加密方案对比与分析

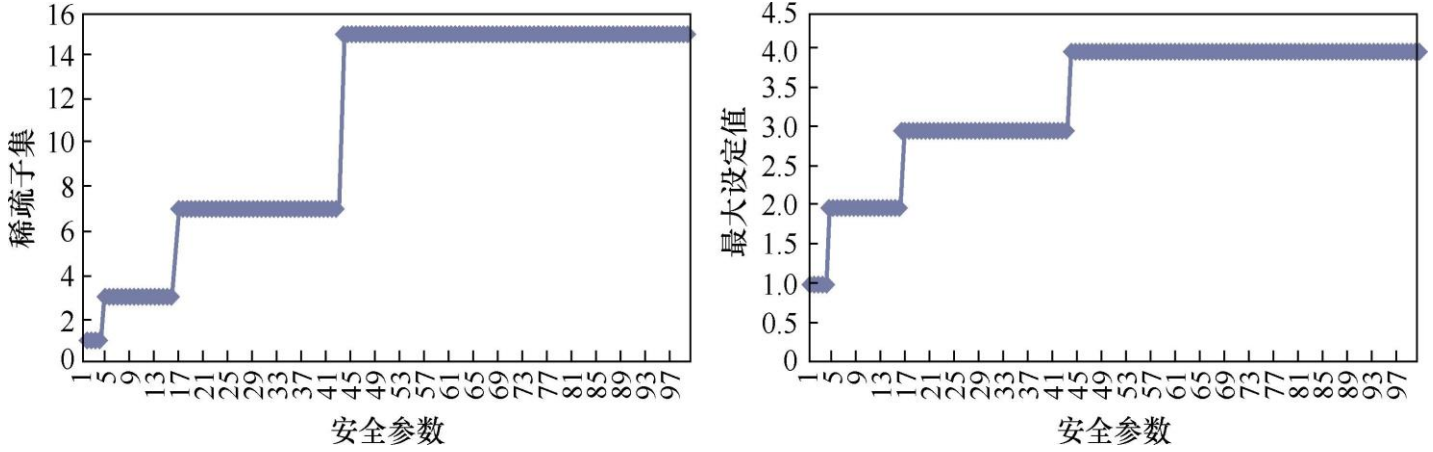


图7.18 稀疏子集、最大设定值与安全参数的关系

小结

云安全概述
虚拟机安全
云存储安全
云数据安全
实践：全同态加密算法

课内复习

1．云计算的安全技术框架包含哪些内容？
2．虚拟化软件栈面临哪些安全威胁？
3．虚拟化软件栈有哪些防御措施？
4．安全云存储系统有哪些关键技术？

课外思考

1．云数据的安全与隐私问题是否会阻止云计算的发展？

2．怎样在云计算的便捷性和云计算的安全问题上进行取舍？

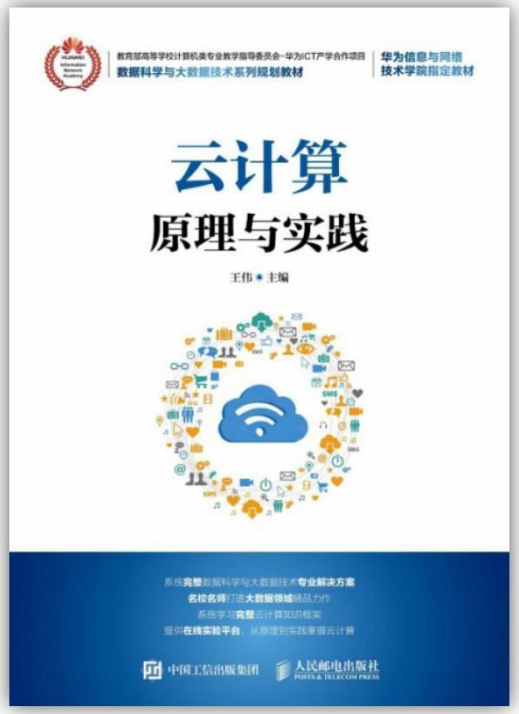
动手实践1

- HElib库是基于C++语言的同态加密算法软件库，能够实现BGV同态加密体制，HElib库的编写基于NTL数学算法库和GMP大数库。
- 任务：通过HElib的项目网站进一步了解并使用HElib。

动手实践2

- FHE-CODE是Gentry的全同态加密方案的一种变体，适 用于64位的Intel处理器。该算法库所依据的算法包括用 于密钥生成算法、加密算法、解密算法、密文刷新算法。
- 任务：通过FHE-CODE的项目网站进一步了解并使用FHE-CODE。

Thanks!


教育部高等学校计算机类专业教学指导委员会-华为ICT产学合作项目
数据科学与大数据技术系列规划教材
华为信息与网络
技术学院指定教材
云计算
原理与实践
王伟 主编
中国工信出版集团
人民邮电出版社
POSTS & TELECOM PRESS