操作系统安全

Operating System Security

[第12次课] 机密计算原理与安全应用

授课教师：游瑞邦

授课时间：2024年5月17日

课程编码：1802030839X4P2002H

授课团队名单：涂碧波、游瑞邦

概 要

- 机密计算概述
- 机密计算原理
- 机密计算安全应用

为何需要机密计算?

数据的三种状态:

Data in Transit

“传输中”状态:
数据正在网络中传输

局域网、互联网……

Data at Rest

“静止中”状态:
存储中的数据

硬盘、U盘……

In use

“使用中”状态:
正在处理的数据

内存、缓存……

加密技术可以用来提供:
- 数据机密性（防止未经授权的访问）
- 数据完整性（防止或检测未经授权的修改）

使用中的数据所面临的安全威胁

“使用中”状态：
正在处理的数据

面临安全威胁

缓冲区溢出

恶意软件和木马

权限提升攻击

零日漏洞

内存泄露

内存侧信道攻击

CPU侧信道攻击

同驻攻击

数据劫持

供应链攻击

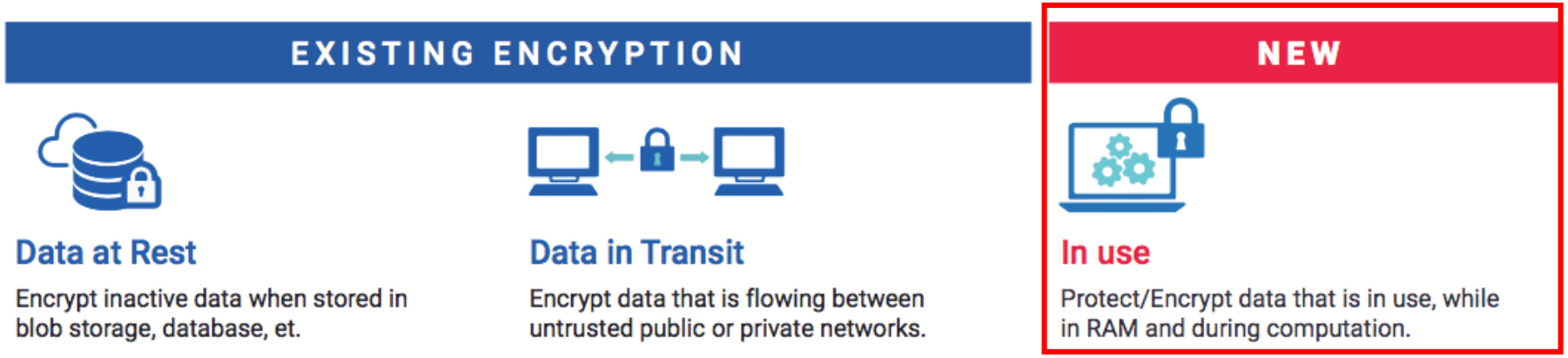
……

安全需求

- 随着越来越多的数据迁移到了云端进行处理，传统的网络安全和物理安全防护机制在防范攻击的能力上越来越有限。已被广泛研究的针对云应用的攻击模式包括虚拟机逃逸、容器逃逸……加剧了使用中的数据安全
- 随着越来越多的数据需要在移动设备、边缘设备和物联网设备上进行存储和处理，实际进行数据处理的地方往往在远端且通常是难以确保其安全性，因此在执行过程中对数据和应用程序提供保护变得越来越重要
- 法律法规

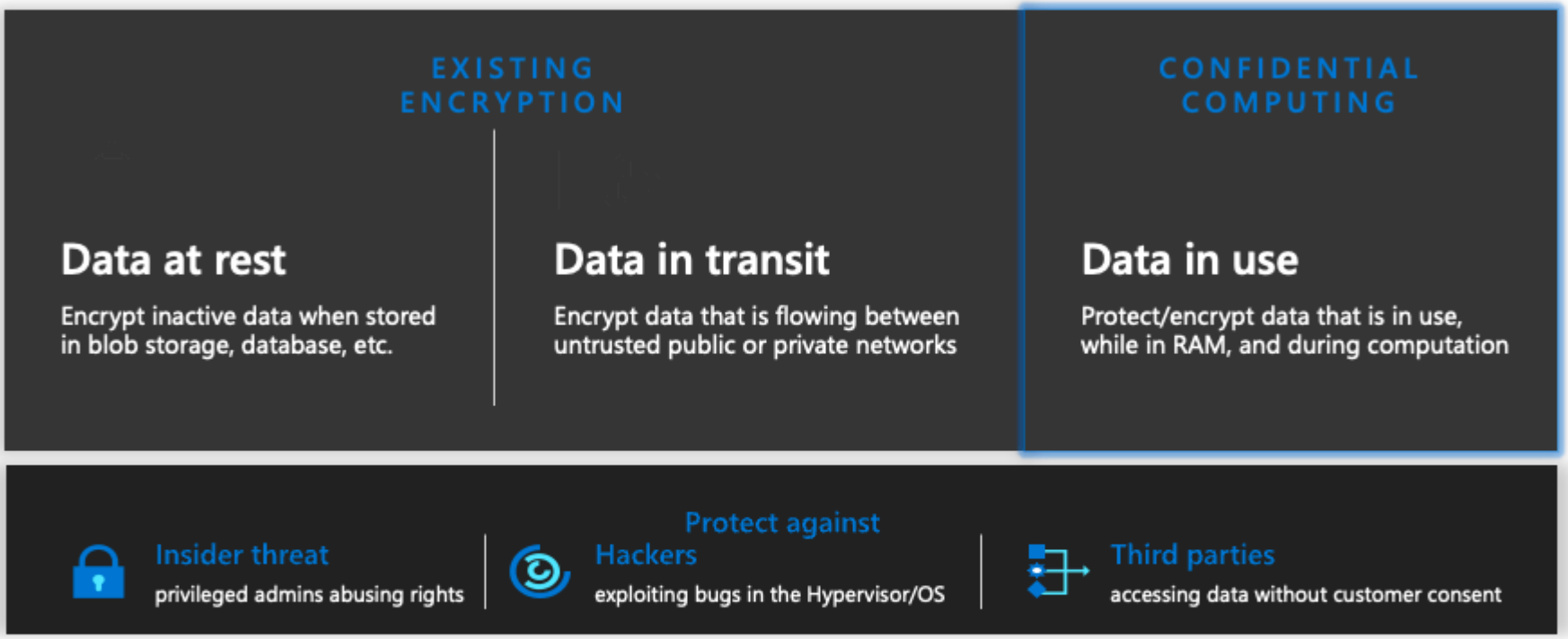
机密计算

- 是通过在基于**硬件的可信执行环境**（Trusted Execution Environment）中执行计算过程的方式，为**使用中的数据提供保护**的计算模式
- 是一种新兴的数据保护方式，旨在**保护数据**在处理时的**隐私**和**安全**
- 通过**确保数据**在**使用过程**中**始终处于加密状态**，来防止未授权访问



机密计算的作用

机密计算通过在**基于硬件**的**经验证**的**受信任执行环境**中执行计算来**保护正在使用的数据**



- **机密计算需要硬件支持的必要性**
  - **整个计算栈中每一层的安全强度都必须至少与它下一层的安全强度一样**

    计算栈的任何一层的安全性都可能被底层的漏洞所规避。上一层的安全依赖于下一层的安全状态，尽可能低的层次上提供更为彻底的安全解决方案，直至硬件层
  - **利用最底层硬件所能提供的安全性，在保持最小信任依赖的条件（TCB）**

    可以将操作系统和设备驱动程序供应商、平台和设备供应商、服务提供商及管理员从需要信任的实体列表中删除，从而减少潜在的风险
  - **机密计算联盟只认可基于硬件的机密计算环境**

    基于软件信任根的TEE排除在外

- **机密计算不限于特定处理器所能提供的可信执行能力，如GPU或网卡也同样能够具备这种可信计算能力**
- **机密计算不局限于数据加密技术，尽管这是机密计算最常用的技术**

机密计算联盟

- 2019 年， CPU 制造商、云提供商和软件公司— 阿里巴巴、AMD、百度、Fortanix、谷歌、IBM/Red Hat®、英特尔、微软、甲骨文、瑞士电信、腾讯和 VMware — 在 Linux 基金会的赞助下成立了机密计算联盟 (CCC-Confidential Computing Consortium)
- CCC 旨在为机密计算定义全行业标准，并促进开源机密计算工具的开发

Premier Members

AMD, arm, Google, HUAWEI, intel, Meta, Microsoft, NVIDIA, Red Hat, TikTok

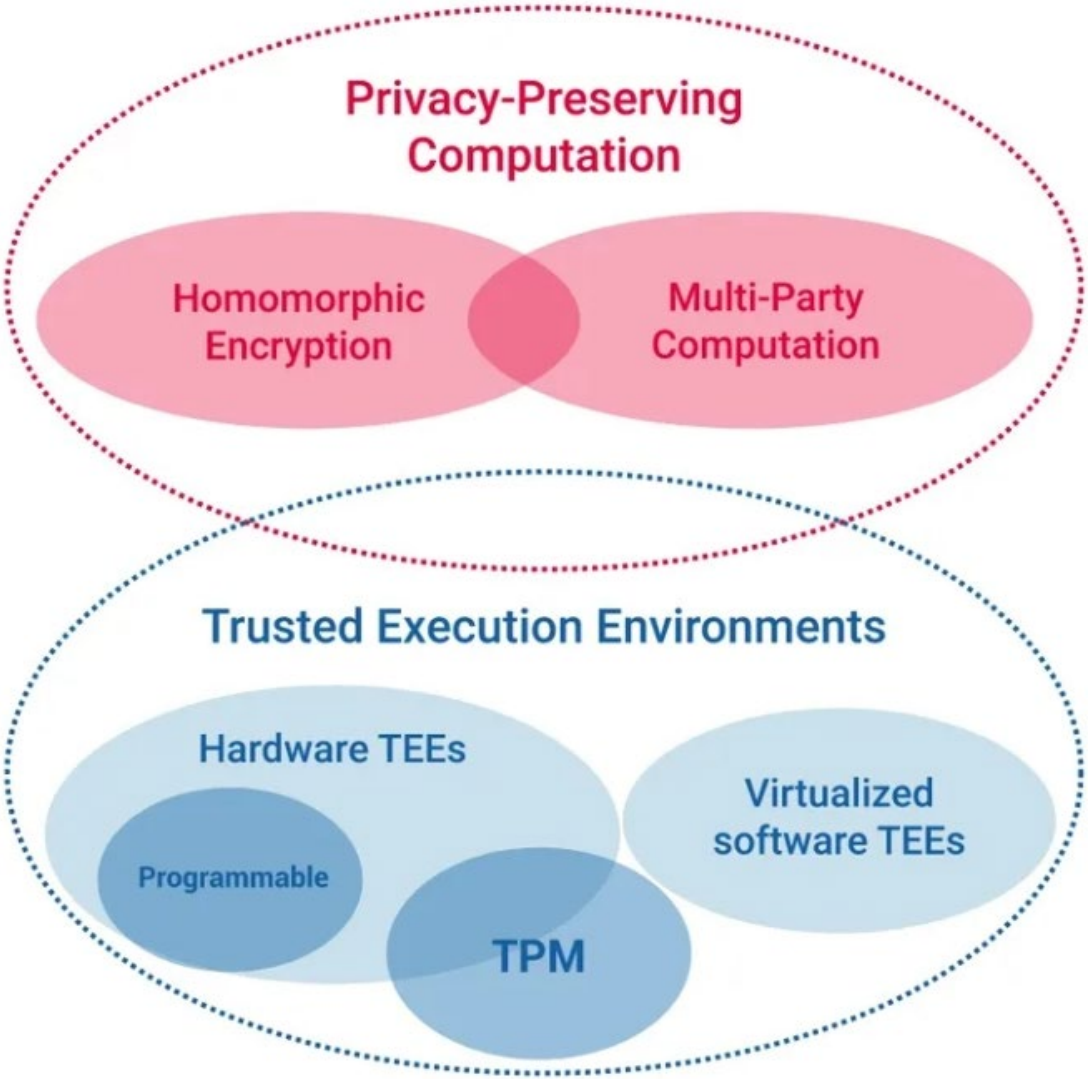
Associate Members

ACURAST, CHMRC, Distributed Ledger Technology, Fraunhofer SIT, Linaro, MIT Connection Science, OASIS FOUNDATION, RISC-V, 上海交通大学 Shanghai Jiao Tong University

General Members

AMPERE., anjuna, 蚂蚁集团 ANT GROUP, BeeKeeperAI, CanaryBit, CANONICAL, cisco, CRYPTOSAT, CYSEC, EDGELESS SYSTEMS, enclaive, ENVEIL, Ethernity CLOUD, FUJITSU, GENXT, hushmesh, iExec, MARVELL, Opaque, Rivos, SAMSUNG, spectro cloud, SUSE, ULTRAVIOLET, vmware

机密计算与隐私计算


Privacy-Preserving Computation
Homomorphic Encryption
Multi-Party Computation
Trusted Execution Environments
Hardware TEEs
Programmable
TPM
Virtualized software TEEs

基于硬件的TEE、可信平台模块（TPM）以及同态加密的安全性比较

	HW TEE	Homomorphic Encryption	TPM
Data integrity	Y	Y (subject to code integrity)	Keys only
Data confidentiality	Y	Y	Keys only
Code integrity	Y	No	Y
Code confidentiality	Y (may require work)	No	Y
Authenticated Launch	Varies	No	No
Programmability	Y	Partial (“circuits”)	No
Attestability	Y	No	Y
Recoverability	Y	No	Y

普通计算、使用典型的基于硬件的TEE计算和同态加密的可扩展性比较

	Native	HW Tee	Homomorphic Encryption
Data size limits	High	Medium	Low
Computation Speed	High	High-Medium	Low
Scale out across machines	Yes	More work	Yes
Ability to combine data across sets (MPC)	Yes	Yes	Very limited

概 要

- 机密计算概述
- **机密计算原理**
- 机密计算安全应用

机密计算的目标

最大限度地降低平台所有者、管理员和攻击者访问TEE内的数据和代码的能力，从而使该路径在执行层面变得不再“经济”或逻辑上可行

机密计算威胁模型的范畴包括：

- 对TEE和TEE环境进行证明，以确保实施了有效和正确的部署
- 将工作负载和数据传输到TEE环境中
- 在TEE实例之外存储与TEE环境相关的数据
- 在TEE环境之间迁移工作负载

- 威胁向量
  - 软件攻击

    包括对主机上的软件和固件的攻击，如操作系统、hypervisor、BIOS、其他软件栈、以及与任何一方有关的工作负载
  - 协议攻击

    包括对与证明以及工作负载和数据传输相关的协议的攻击。如任何可能危及证明TEE实例的攻击危及到工作负载或数据的安全，在工作负载和/或数据的配置或部署过程中发生的漏洞等
  - 密码攻击

    密码和算法中的漏洞，包括数学上的突破、计算能力的提升和新的计算方法，如量子计算
  - 基本的物理攻击

    包括冷DRAM提取（cold DRAM extraction）、总线和缓存监听以及将攻击设备插入端口的攻击，比如PCIe、Firewire、USB
  - 基本的上游供应链攻击

    包括软硬件植入后门/漏洞等，如通过添加调试端口全局性的危害TEE组件的攻击

- **复杂的物理攻击**

  通常需要长期和/或侵入性地访问硬件，包括芯片scraping技术和电子显微镜probes。

- **上游硬件供应链攻击**

  这类攻击包括对CPU的攻击，比如发生在芯片制造阶段和密钥注入/生成阶段的攻击，但不包括针对主机系统中不直接提供TEE功能的组件的攻击。

- **可用性攻击**

  在当前基于硬件的TEE的安全威胁模型中，是不解决可用性的，如DoS或DDoS

这三项属于机密计算范畴之外的威胁向量

- 可信执行环境（TEE）

  机密计算联盟将可信执行环境定义为一个能够为以下三个属性提供一定程度保证的环境：

  - 数据机密性：未经授权的实体不能窥探到在TEE内使用的数据
  - 数据完整性：未经授权的实体不能添加、删除或篡改在TEE内使用的数据
  - 代码完整性：未经授权的实体不能添加、删除或篡改在TEE中执行的代码

未经授权的实体

在机密计算的上下文中，未经授权的实体包括主机上的**应用程序**、**主机操作系统**和**Hypervisor**、**系统管理员**、**服务提供商**、**基础设施所有者**或**任何能够物理接触硬件的人**

核心组件

机密计算的实现主要依赖于以下几个核心组件：

- **可信执行环境（TEE）/安全执行环境（SEE）：**

  TEE或SEE提供一个隔离的执行空间，保护其中的代码和数据免受外部环境的访问。这种环境通常由硬件支持，并可防止来自操作系统、虚拟机监视器或其他应用程序的访问

- **硬件信任锚点：**

  硬件信任锚点是一种可信的物理组件，用于确保只有经过授权的代码可以在TEE中执行。这包括安全模块（如TPM）或特定的CPU指令集，它们用于启动安全启动序列和验证TEE的完整性

- **加密内存：**

  机密计算使用加密内存技术保护物理内存中的数据。即使有人能够直接访问物理内存，未经授权的访问者也无法读取或修改数据

- **远程证明：**

  远程证明是一种机制，TEE可以通过它生成一个证明，表明它是按照特定的安全策略配置和启动的。这允许第三方验证器（如数据所有者）确信其数据只会在一个安全且信任的环境中被处理

- **数据隔离：**

  数据在进入TEE之前被加密，在离开之前重新加密，只有在TEE内部才被解密处理。这确保了数据的安全性和隐私性，即使在跨越不受信任网络或环境时也是如此

- 证明（Attestation）
  - 证明是一个过程。通过该过程，作为验证者的一方将评估潜在不可信的另一方，即证明者的可信度
  - 证明的目标是通过获得真实、准确和及时的关于证明者的软件和数据状态的报告，使验证者能够相信证明者

基于硬件的证明

基于硬件的证明方案依赖于在安全环境中可信的硬件组件和相关固件所执行的证明步骤：

- 在验证者和证明者之间建立一条安全的通信信道
- 验证者生成挑战值并将其发送给证明者
- 证明者通过将挑战值发送到其信任的硬件组件并请求返回包含了其软件和数据状态的证据
- 可信的硬件组件在证明平台上收集证据并对证明数据和挑战值进行签名
- 证明者将经过签名的证据返回给验证者
- 验证者验证签名，并根据鉴定策略来鉴定证据，比如将已证明的平台状态与一组可信参考值进行比较，并验证已签名的证据中是否包含了之前提供的挑战值，以便说服验证者相信证据是新生成的

- 机密计算的匿名性
  - 在基于硬件的证明方案中，能够通过可信的硬件组件相关联的密钥来唯一标识该硬件组件。攻击者可能会利用这一点来监控特定可信的硬件组件的活动
  - 实现匿名的一种做法是采用直接匿名认证（DAA-Direct Anonymous Authentication）方案。该方案利用先进的加密原语，如零知识证明和组签名，来应对这一隐私挑战。DAA方案可以使用多种密码技术实现，包括RSA、椭圆曲线密码（ECC）、基于配对的密码（PBC）或基于晶格的密码（LBC）

机密计算的TCB恢复

- **Trusted Computing Base（TCB）恢复指的是在发现了TEE的TCB中存在可以修复的缺陷后进行修复，并最终能够重新建立对TEE的信任的过程**

  如经过签名的证据与最新的参考值不匹配，则证明过程会让验证者得出证明者是不完全可信的结论，这时就需要进行TCB修复

- **TEE的TCB（通常包括不可变部分和可变部分）提供了能够创建用在证明过程中的证据的功能**

  如在TCB中发现了缺陷，则证明过程本身可能被欺骗或破坏。TEE实现可以利用特殊技术，让不可变部分来保障对可变部分的更新过程。验证者可以通过识别更新后新创建的任何证据，证实确实已经进行过TCB更新，而且不会被以前有缺陷的实现所欺骗

- 机密计算依赖的硬件技术——硬件可信执行环境
  - ARM TrustZone
  - ARM CCA
  - Intel SGX
  - Intel TDX
  - AMD SME SEV
  - AMD SEV-SNP

- 机密计算主要使用方式
  - 应用SDK方式
    - 开发人员负责将其应用程序的代码划分为可信组件和不可信任组件
    - TEE提供商将SDK抽象成了一个通用的编程模型，同时提供了跨硬件TEE支持的可移植性
  - 运行时部署系统
    - 最小化了将典型的应用程序工作负载转换为可以运行在TEE中的工作负载所需的工作量
    - 所开发的应用具有跨TEE可移植性的，甚至可支持将未修改的应用程序直接部署到TEE中
  - 容器化运行
    - 机密容器部署方式

概要

- 机密计算概述
- 机密计算原理
- **机密计算安全应用**

机密计算安全应用

- 核心场景
  - 存储和处理密钥、秘密信息、凭证以及令牌
  - 密钥、秘密信息、凭证和令牌是保护敏感数据的关键信息资产
  - 当地国家安全标准的本地硬件安全模块（HSM）对这些关键信息资产进行存储和处理
  - 密钥管理应用程序可以在基于硬件的安全TEE中存储和处理密钥、秘密信息和令牌，并提供数据机密性、数据完整性和代码完整性，以实现与传统HSM同等的安全性

主要应用场景

包括公有云、多方计算、区块链、个人移动和计算设备、边缘和IoT、POS机 / 支付

Use case	Sub-use case	Everest Group definition
Privacy and security	Key management system	Confidential computing used to secure systems that deal with the generation, exchange, storage, and use of cryptographic keys used across the enterprise IT landscape
	Application security on public cloud	Confidential computing used in typical cloud scenarios such as containers or microservices applications to prevent the compromise of data from malicious actors and proactively improve an application's security and privacy posture
Blockchain	Meaningful Proof of Work (PoW)	PoW is a cryptographic concept in blockchain that helps one party prove to others in a blockchain network that a certain amount of a specific computational effort has been expended; this effort can then be verified through minimal effort on the blockchain. Meaningful proof of work is a technique of PoW where the computational effort required is utilized for productive purposes
	Blockchain data privacy	Confidential computing used to secure data being written or processed on the blockchain network to add an additional element of privacy to the system
Multi-party computing	Private data sharing	Utilization of confidential computing by an entity such as a hospital or bank, to share sensitive data, such as patient or financial data with third parties such as pharmaceutical companies, insurers, or credit rating agencies
	Multi-party analytics	Use of confidential computing by multiple entities aggregating their proprietary data and collaboratively analyzing it to gain new insights
	Privacy-preserving AI/ML modeling	Use of confidential computing to secure data during the modeling and training of AI; further confidential computing is used to secure modeling in decentralized training mechanisms such as federated learning
IoT and Edge	Trusted command and control	Confidential computing used to protect critical infrastructure connected to the internet to prevent malicious entities from accessing, controlling, or manipulating key devices, sensors, or systems
	Secure data and IP	Use of confidential computing to secure intellectual property and data generated or utilized in edge and IOT devices from malicious elements
Personal computing devices	Personalized recommendations	Confidential computing used on personal computing devices to analyze data and build models on the device to reduce the need for off-device or cloud processing

机密计算的思考

- “可见不可用”
- “可用不可识”
- 隐私计算与机密计算
- 机密计算与机密容器


Q&A