安全+
SECURITY
2024/12 总第 063
技术版 与安全人士分享技术心得 Share technique experience with security professionals

★ 本期焦点

基于虚拟机级TEE技术的应用与性能分析研究

构建精益安全体系：浅析4个攻击面管理技术

网络安全政策导读（2024年9—12月）

NSFOCUS

本期看点 HEADLINES

主办：绿盟科技
策划：《安全+》编委会
地址：北京市海淀区北洼路4号院绿盟科技园
邮编：100089
电话：(010)6843 8880-5462
传真：(010)6872 8708
网址：www.nsfocus.com

2024/12 总第 063

欢迎您来信nsmagazine@nsfocus.com 与我们交流，分享您的建议和评论。（《安全+》部分图片来源于网络）

安全+ SECURITY

目录 CONTENTS

随着人工智能和数字化技术的加速推进，网络安全领域正在迎来全新的发展阶段。2024年，“AI+安全”逐步从技术愿景走向实际落地，各行各业在AI赋能下展现出新的活力，同时也面临更多样化的安全挑战。从国家政策的顶层设计到产业实践的创新探索，网络安全已经成为数字化转型不可或缺的关键基石。

国务院印发的《关于推进新型工业化的指导意见》强调，以“人工智能+”为抓手，加快产业现代化进程。在此背景下，网络安全产业既是产业升级的重要保障，也在不断挖掘自身价值。安全技术的升级，不再只是“防御”的延续，而是以赋能生产力为核心，为数字经济的高质量发展提供底层支持。

在这一背景下，绿盟科技持续深化“智慧安全”理念，围绕AI赋能、数据安全、云原生安全等核心方向推出了多项创新成果，推出的风云卫AI安全能力平台，以大模型为驱动，构建了新一代智能安全运营体系，实现了从防护到主动安全的升级。我们还在机密计算、攻击面管理、蓝军对抗等领域开展深入研究，力求通过技术创新解决复杂场景中的实际问题。

本期《安全+》聚焦“技术与实践并行”，包括虚拟机级TEE技术的性能分析、CAASM框架的落地思考，以及攻击面管理的应用等前沿课题。在国家政策支持和技术快速演进的双重驱动下，网络安全不再只是“防御”，更是“赋能”。如何以精益安全为核心，构建支撑未来数字经济的信任体系，将是我们持续探讨的重要命题。

网络安全的未来，不仅是一场技术的较量，更是一场生态的协同。在此，我们希望通过《安全+》，与更多伙伴共同探索前沿技术、分享落地经验，以智慧安全赋能数字未来。

叶晓虎

基于虚拟机级TEE技术的应用与性能分析研究

绿盟科技 创新研究院 陈佛忠

摘要：本文从计算类型、数据源和数据管理策略这三个关键维度出发，系统地分类了隐私计算的应用场景，并进行了虚拟机级 TEE 技术适用性的定性分析。此外，通过精确的实验方法，本研究对虚拟机级 TEE 技术的性能进行了定量评估，并将其与其他隐私计算技术进行了比较。

关键词：可信执行环境 隐私计算 数据安全 数据共享

1. 隐私计算概述及相关政策动态

在当今信息时代，数据成为无处不在的宝贵资源。从金融机构到医疗保健领域，从科学研究到政府机构，巨大的数据集用于支持决策制定、创新发展和提供各种服务。然而，数据加工生产和交易流通中信息的共享、收集、发布、分析与利用等操作会直接或间接地泄露用户隐私，给用户带来极大的威胁和困扰，这引发了关于如何更好地保护敏感信息的担忧 [1]。

正是在这个背景下，隐私计算技术崭露头角。隐私计算，也称为隐私增强计算，是一种在保障隐私的前提下实现数据价值的技术体系。它通过多方安全计算、联邦学习和可信执行环境等技术，构建用于数据保护的特定技术方案。隐私计算允许在加密或非透明状态下进行数据处理，实现多方数据共享，同时确保所有参与方的隐私得到保护 [2][3]。

在 2023 年里，国内多个部门相继推出了关于隐私计算的政策，从上到下对隐私计算技术落地应用的热潮日益高涨。例如，2023 年 7 月 24 日央行发布《中国人民银行业务领域数据安全管理办法（征求意见稿）》，指出数据处理者采用隐私计算等技术促进数据融合创新应用时应当确认原始数据未离开自身控制范围，且多个数据提供行为关联后，暴露约定范围外信息的风险可控 [4]。2023 年 12 月 19 日工信部发布《工业领域数据安全标准体系建设指南（2023 版）》，提出将多方安全计算、联邦学习等作为数据共享安全技术产品标准重点建设方向，将数据脱敏、可信执行环境等作为数据安全防护技术产品标准重点建设方向 [5]。2023 年 12 月 31 日国家数据局等 17 个部门共同印发了《“数据要素 ×”三年行动计划（2024—2026 年）》，在该计划中明确提到需要打造安全可信流通环境，深化数据空间、隐私计算、联邦学习、区块链、数据沙箱等技术应用 [6]。

2. 虚拟机级 TEE 技术简介及分析

可信执行环境（trusted execution environment，TEE）是隐私计算技术之一。这一技术的核心思想是通过 CPU 级别的隔离和保护机制，创建一个与操作系统或其他应用程序隔离的安全区域，以确保数据的安全性。通常情况下，计算机中内存与磁盘数据是以明文状态存在的，这会使它们容易受到未经授权的软件或管理员的查看或篡改。TEE 提供了一个受保护的环境，确保只有经过授权的计算进程能够访问和操作数据，而不会受到外部干扰或泄露的影响 [7]。

在 TEE 技术中，不同的实施方式定义了哪些元素有权访问机密数据，当前可商用技术主要分为两类：虚拟机级（VM-TEE）和进程级 TEE 技术。VM-TEE，如 AMD SEV，仅允许在基础设施上运行的特

定虚拟机内的元素访问敏感数据。而进程级 TEE 技术，如 Intel SGX，则仅授权特定的软件应用程序或进程来访问这些数据 [8]。

相比于算法类隐私计算技术，如联邦学习（FL）和多方安全计算（MPC），TEE 技术在计算性能上展现了显著的优势，本文将在第四部分通过模拟实验进行详细的定量对比分析。在实际应用部署方面，VM-TEE 能够直接将计算程序导入并进行安全调度，而进程级 TEE 技术则需要对计算程序进行额外地修改以确保其适配性。

3. 应用场景分析

隐私计算场景通常与业务需求紧密结合，因此具有一定的复杂性。本文将根据市场上的普遍需求，主要从计算类型、数据源和数据管理策略三个维度进行详细分析。计算类型主要包括 AI 计算和任意计算两种；数据源分为单方数据源和多方数据源；数据管理策略则分为数据本地化和数据密态出域两种情况。AI 计算专注于通过训练 AI 模型进行业务分析；而任意计算涵盖了包括 BI 分析和 AI 分析在内的各类数值计算；单方数据源指的是除查询条件外，数据仅来源于一个单一方；而多方数据源则涉及两个或更多的数据方；数据本地化是指数据无论是明文还是密文形式，都不允许出本方的网络域，通常出现在安全性要求较高的机构；数据密态出域是指仅允许数据以加密的形式共享出去。通过这些维度的不同组合，我们可以划分出如表 1 所示的八种应用场景。

表 1 应用场景分类

	场景 1	场景 2	场景 3	场景 4	场景 5	场景 6	场景 7	场景 8
计算类型	AI 计算				任意计算			
数据源	单方		多方		单方	多方	单方	多方
数据管理策略	数据本地化	数据密态出域	数据本地化	数据密态出域	数据本地化		数据密态出域	

对于场景 1 和场景 2，在 AI 计算中，由于数据集仅限于单一方，通常不存在数据共享和交互的需求。因此，这些场景不在本次讨论的范围之内。

3.1 多方数据源的 AI 计算场景（场景 3、场景 4）

以医院或医疗研究团队利用 AI 技术分析特定疾病（如心脏病、糖尿病等）的发病可能性为例。通常，单一机构仅能获取有限的真实病例数据，因此为了提升预测的准确性，必须依赖大量数据进行模型训练，这往往需要与外部团队合作。场景 3 和场景 4 的主要区别在于数据管理策略：在场景 3 中，所有参与方的数据必须保留在本地；而在场景 4 中，可以允许数据在保持密态的情况下跨域传输。

基于场景 3 的需求，目前市面上的常见解决方案是 FL。FL 能够在确保合作各方数据不离开本地的条件下，实现分布式 AI 建模和预测。然而，这种方案尚未广泛实施，主要障碍包括较高的建设成本（各方都需部署系统）、运行成本（各方都需提供计算资源）和沟通成本。在这种场景下，TEE 技术目前并不适用。

在现实生活中，如场景 4 的集中式 AI 建模更为常见，这通常由较大的机构或主管部门负责推动。例如，卫生健康委员会可能会集中管辖区内多家医院的数据进行本地 AI 模型分析。但是，传统的机器学习和深度学习技术无法充分保障数据在运行态的安全性，运维人员有可能通过 dump 内存的方式获取到明文数据。在这种场景下，利用 VM-TEE，可以实现数据在运行态时的机密性，从而有效降低数据从内部泄露的风险。尽管这类场景非常常见，但是目前几乎没有隐私计算落地的案例。

3.2 数据本地化的任意计算场景（场景 5、场景 6）

数据本地化的单方任意计算在真实生活中非常常见，例如查某人的征信情况、是否有犯罪前科、是否有精神疾病等。这类业务场景相对简单，但涉及的数据极其敏感，通常数据会被储存在本地，且严禁以任何形式泄露。在这些场景中，使用 VM-TEE 技术可以同时保护查询条件对数据源方的隐蔽性和数据源的信息对查询方的不可见性。从技术角度看，MPC 也能达到同样的效果，但它会显著增加计算与通信的成本，并且在复杂的计算场景中还需增加额外的编码工作量。

数据本地化的多方任意计算实现起来会比较困难，以查询个人在各个银行的贷款情况为例，这需要在业务流程中说服每个银行协同工作，并且需要在各方本地部署软件或硬件设备。从技术角度来看，虽然 MPC 技术可以在数据不离开本地的前提下进行联合统计分析，但其计算和通信成本比场景 5 更高。对于 TEE 技术，各方可以在本地进行计算，并汇总这些结果以得到最终输出。在此过程中，只有加密的计算结果会离开本地环境。因此，参与方可以根据计算结果的敏感度来决定是否支持此类操作。

3.3 数据密态出域的任意计算场景（场景 7、场景 8）

场景 7 以大数据局为例，大数据局会汇集多个政府机构的数据来进行分析，其中有些简单的场景仅涉及单方数据源（如统计本地的出生率、死亡率）。场景 8 以公安部门为例，其在疫情期间可能需要汇总多个社区的感染情况，以便对感染者实施重点防控措施，有效减缓疫情的传播。

在这些场景中，无论数据之前是否经过加密传输或存储，其在运行态时均为明文，运维人员通过 dump 内存可以获取到明文数据。利用 TEE，可以实现数据在运行态时的机密性，从而有效降低数据从内部泄露的风险。对于 VM-TEE 来说，计算程序无须额外改造，可以直接将其容器化后导入 TEE 中，这个技术优势大大减少了计算程序改造的成本。尽管这类场景非常常见，但是目前几乎没有隐私计算落地的案例。

3.4 小结

场景 4、场景 7 和场景 8 允许数据以加密形式共享，这在现实中非常普遍。然而，这些场景很少采用隐私计算技术来确保安全。根据 Verizon 最新发布的《2024 年数据泄露调查报告：漏洞利用增长威胁网络安全》，68% 的数据泄露事件源于内部人为因素 [9]。利用 TEE 技术可以有效减少内部数据泄露的风险，预计未来将有更多隐私计算技术应用于此类场景。对于场景 3、场景 5、场景 6，即数据不离开本地的情况，虽然已有少数隐私计算案例落地，但尚未形成规模化应用。

表 2 适用性分析

	场景 3	场景 4	场景 5	场景 6	场景 7	场景 8
隐私计算的应用案例	少数 FL 的案例	几乎无	少数 TEE、MPC 的案例		几乎无	
进程级 TEE	不适用	一般适用（需要额外编码程序）				
VM-TEE	不适用	适用				

4. 性能分析

由于 FL 技术无法完成任意计算，本文仅实验对比 MPC 与 VM-TEE 技术的性能差异。不失一般性，本文对场景 6 选取两方集合求交这一实例进行测试，即两个参与方各拥有一个集合，双方希望在不暴露集合内容的情况下获得两方集合的交集；集合的数量是影响计算时间的核心因素。

对比实验考虑到 MPC 技术（基于 KKRT 技术 [10]）、VM-TEE 技术加密（基于海光 CSV 与 luks 磁盘加密）与无加密的明文三种情况。实验机器配置为 8 核 16G 物理机或加密虚拟机，机器处于同一局域网中，网络时延低于 1ms；总体时间消耗如表 3 所示，其中 VM-TEE 与明文包含数据传输汇聚与实际计算两阶段。

表 3 不同技术下两方集合求交时间消耗 (单位为 s)

每方集合内元素数量	MPC	VM-TEE			明文		
	总耗时	总耗时	数据传输汇聚时间	实际计算时间	总耗时	数据传输汇聚时间	实际计算时间
10^4（万级）	0.76	0.29	0.28	0.01	0.29	0.28	0.01
10^5（十万级）	14.82	0.44	0.28	0.16	0.43	0.28	0.15
10^6（百万级）	164.78	2.77	1.14	1.63	1.92	0.41	1.51
10^7（千万级）	1718.17	19.39	4.18	15.21	15.95	1.3	14.65

实验数据显示，相比于明文，采用 MPC 方案时间消耗会提升两个数量级；而 VM-TEE 方案受网络虚拟化影响，大文件传输时延至多提升 220%，实际计算时延至多仅提升 3.8%。

5. 总结

随着数据要素的快速发展，隐私计算应用的趋势也在不断加快。在应用场景分析中，本文从计算类型、数据源和数据管理策略三个关键维度，对隐私计算的应用场景进行了详细分类。在性能分析部分，本文通过实验方法比较了 VM-TEE、MPC 和明文三种情况下的计算时间。结果显示，VM-TEE 技术在性能上具有显著优势，尤其是在处理大量数据时，其时间消耗明显低于 MPC 技术。实验数据表明，VM-TEE 技术在网络虚拟化影响下，大文件传输时延至多提升 220%，实际计算时延至多仅提升 3.8%，这充分证明了 VM-TEE 技术在隐私计算中的高效性和实用性。

综上所述，本文系统地探讨了隐私计算的应用场景，并通过定量实验验证了虚拟机级 TEE 技术的性能。研究结果表明，VM-TEE 技术不仅能有效保护数据在运行态的安全性，还能显著降低计算和通信成本。随着隐私计算技术的不断发展和政策的推动，VM-TEE 技术在未来数据安全领域中将发挥越来越重要的作用，为数据的安全流通和价值创造提供坚实的保障。

参考文献

[1] 李凤华，李晖，贾焰，等 . 隐私计算研究范畴及发展趋势 [J]. 通信学报，2016, 37(4):1—11.

[2] AGRAWAL R, SRIKANT R. Privaey-preserving data mining[C]// ACM SIGMOD Reeord. c2000:439-450.

[3] AGGARWAL C C, YU P S. A general survey of privacy-preserving data mining models and algorithms[M]. Privacy-Preserving Data Mining. Springer US, 2008:11-52.

[4] 中国人民银行 . 关于《中国人民银行业务领域数据安全管理办法（征求意见稿）》公开征求意见的通知 [Z]，2023.

[5] 工业和信息化部，国家标准化管理委员会 . 关于印发《工业领域数据安全标准体系建设指南（2023 版）》的通知 [Z]，2023.

[6] 国家数据局等十七部门 . 关于印发《“数据要素 ×”三年行动计划（2024—2026 年）》的通知 [Z], 2023.

[7] A. Akram, V. Akella, S. Peisert and J. Lowe-Power. SoK: Limitations of Confidential Computing via TEEs for High-Performance Compute Systems [J]. 2022 IEEE International Symposium on Secure and Private Execution Environment Design (SEED), Storrs, CT, USA, 2022, pp. 121-132.

[8] Confidential computing consortium. Common Terminology for Confidential Computing[R], 2022.

[9] Verizon.《2024 年数据泄露调查报告：漏洞利用增长威胁网络安全》[R]，2024.

[10] Vladimir Kolesnikov, Ranjit Kumaresan, Mike Rosulek, and Ni Trieu. 2016. Efficient Batched Oblivious PRF with Applications to Private Set Intersection [J]. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16). Association for Computing Machinery, New York, NY, USA, 818–829.

机密计算联盟开源项目之机密计算认证框架

绿盟科技 创新研究院 王拓

摘要：远程证明是机密计算的核心能力之一，其旨在向远端用户证明，他们的工作负载是否真的运行在机密环境中。然而机密计算所依托的可信执行环境（TEE）技术种类繁多，各 TEE 技术供应商所提供的远程证明架构、接口等各不相同，这为机密计算应用开发者增加了适配难度。对此，本文介绍了机密计算联盟的开源项目“机密计算认证框架”，分析了其如何整合不同架构下的 TEE 远程证明能力，帮助开发者低成本实现统一的跨平台机密计算远程证明功能。

关键词：机密计算 可信执行环境 远程证明

1. 概述

机密计算联盟（Confidential Computing Consortium，CCC）[1] 是 linux 基金会的一个项目社区，致力于定义和加速机密计算的采用。联盟主要成员包括 arm、谷歌、华为、英特尔、微软等巨头企业。目前 CCC 拥有 11 个项目，本文主要对其中的“机密计算认证框架”（Certifier Framework for Confidential Computing）这一项目进行介绍，该项目由 VMWare 公司启动，旨在通过提供简单的客户端信任管理，简化和统一了对多供应商机密计算平台上有关可信认证的编程和操作支持。若想了解更多可访问该项目 github 仓库 [2]。

2. 为什么需要机密计算认证

机密计算技术可有效保护运行在不安全环境中程序与数据的机密性、完整性和策略合规性，程序的拥有者可以将程序部署在远端不受信环境上，通过机密计算技术来保证自身程序与数据安全。一般来讲，机密计算要求平台具有以下能力：

（1）隔离能力：将程序与其他软件隔离；

（2）度量能力：为程序提供不可伪造的度量信息；

（3）机密存储能力：在隔离的情况下，仅被度量的程序可以获取机密信息；

（4）证明能力：平台可以对隔离程序的行为提供签名证明，证明该程序运行在隔离环境中。

其中证明能力是为机密计算提供安全性背书与证明的重要手段，程序拥有者可以通过机密计算平台所提供证据验证程序所在平台信息、程序完整性以及程序执行策略完整性等信息，来确保自己的程序运行符合预期，没有受到篡改，并运行在隔离环境中。

然而，机密计算的隔离性大多依赖于可信执行环境技术所提供的硬件隔离，如 AMD 的 SEV 与英特尔的 SGX、TDX，ARM 机密计算架构、RSIC-V Keystone 等。各家厂商所提供的软硬件架构、开发方式、迁移方式、证明接口均各不相同。开发者需要针对不同硬件平台做大量有针对性的开发工作，自定义配置不同的证明与信任策略。为了简化与统一上述过程，机密计算认证框架应运而生。

机密技术认证框架主要目标在于减少有关认证部分的程序构建难度，可以支持不同架构的机密技术平台，通过提供简单的 API 来取代原有的大量开发工作，提供机密计算原语的统一接口，包括度量、机密存储、验证等能力，如图 1 所示。

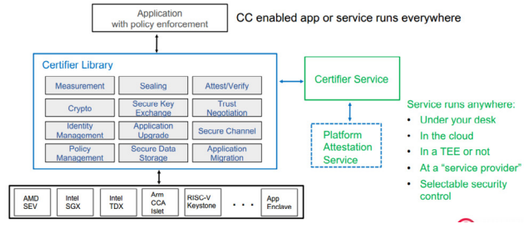


图 1 机密计算认证框架架构

3. 机密计算认证框架介绍

机密计算认证框架由称为认证 API（Certifier API）的客户端和称为认证服务（Certifier Service）的策略评估服务器组成。

- 认证API：提供简单的客户端信任管理，包括证明评估、安全存储、平台初始化、机密共享、安全通道及其他多种服务；
- 认证服务：支持可扩展的策略驱动型信任管理，包括证明评估、应用程序升级和其他机密计算信任服务。

其中，认证 API 的核心为一个认证库（Certifier Library），该认证库将机密计算中远程证明能力以及对可信环境中身份、密钥、证书的管理能力进行封装，抽象为多个接口暴露给开发者，开发者在开发可信应用时无须关注可信证明部分，用简单的接口完成对程序身份、密钥、策略等内容的证明信息生成与管理，其核心思想主要为三个概念：

（1）安全域：可信环境中所有程序代码都有一个与之关联的公钥来标识其身份；

（2）认证：对安全域中所有属性进行认证，生成安全证书；

（3）信任与策略：程序运行策略通过 API 进行管理，而不是硬编码至应用程序中。

基于上述思想，可信环境中所有程序身份、策略、度量、公钥均可由可信环境签名背书，而各个程序或策略等可以使用与之身份对应的私钥所涉及的其他部分（如机密信息、运行策略等）进行签名背书，最终形成信任链，完成对可信环境中所有内容的度量验证。由于认证库对上述功能进行了封装，开发者可以根据接口在不同平台上统一部署，利用接口管理与连接安全信道进行通信，做到跨平台度量认证、身份验证、协同计算等。

而认证服务则可以视为一个中心化的安全域管理服务，其本身由策略驱动，可以动态扩展，其主要功能包括对指定程序升级而不影响其他程序、引入新组件而不影响旧组件、实施安全域范围的策略、促进域内共享与迁移等。简单来说，认证服务可以说是对所有使用了认证 API 的可信程序的统一密钥分发、可信证明验证、策略评估的证书机构，提供类似于 CA 的功能，如图 2 所示，注意如果是 SEV 或 TDX 这种加密虚拟化技术，则不存在 SDK。

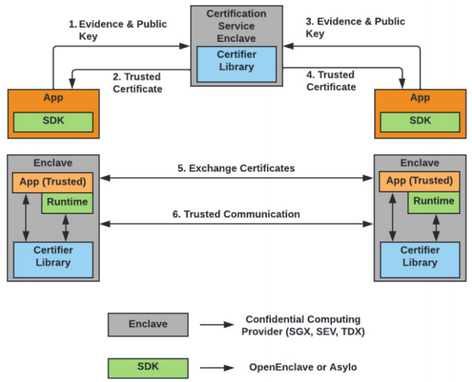


图 2 认证库与认证服务的关系

4. 总结与展望

可以看出该项目主要目的是为机密计算提供跨平台的、开发者友好的统一认证服务架构，可以更容易地为不同平台的机密计算程序提供统一证明能力，证明范围更为丰富与完善，这种统一的架构有助于多个机密计算参与方，在跨平台情况下完成更便捷的信任关系建立与协调计算。该项目在 2023 年 10 月 5 日的报告中指出，目前该项目支持 SGX、AMD-SEV-SNP、ARM CCA、RISC-V Keystone 以及未来正式推出的 TDX 这些机密计算硬件平台。此外，该项目目前也在致力于对 GPU 机密计算环境的相关证明（目前只提到了针对英伟达的 H100），但在程序开发方面，当前示例程序均为 C++ 程序，项目宣传即将完成对 python 的支持，其他语言的支持情况暂不可知。总体来说，该项目凭借其对跨平台机密计算的支持与完善丰富的证明能力让人眼前一亮，期待该项目在未来能够更加丰富与完善，进一步降低机密计算平台的开发成本。

参考文献

[1]https://confidentialcomputing.io/

[2]https://github.com/ccc-certifier-framework/certifier-framework-for-confidential-computing

从传统企业资产管理到CAASM

绿盟科技 创新研究院 张承万

摘要：企业资产管理是信息安全的基础，在企业安全运营中发挥着至关重要的作用。然而，传统 IT 资产管理因分散异构模式面临影子 IT 资产增多、云资产管理不足、数据孤岛等问题。CAASM（网络资产攻击面管理）通过整合多源数据，提供统一资产视图及弱点分析，逐步成为解决企业资产管理痛点的重要工具。以 Axonius 和 JupiterOne 为代表的厂商，分别通过强大的数据接入能力与知识图谱建模优化资产管理。未来 CAASM 需结合 AI 技术，提升自动化威胁检测与风险管理能力，为企业提供更精准的安全防护方案。

关键词：攻击面管理 资产管理 安全运营 内网测绘

1. 资产管理简介

企业资产管理是一个重要且长期的话题。根据 NIST 的定义 [1]，任何对企业有价值的事物都被视作资产，它可以是有形的（如建筑、设备）或无形的（如数据、知识产权）。在信息技术领域，资产管理通常指 IT 资产管理（图 1 为 IT 资产的常见分类），其管理对象涵盖支持业务目标的所有 IT 资产，管理内容包括采购、部署、监控、维护、更新和处置等环节。有效的资产管理对资产的可见性、安全性、合规性和成本控制至关重要。

图 1 IT 资产分类

在实际业务中，由于 IT 资产的复杂性，往往会有多个团队参与资产管理相关工作。如图 2 所示，IT 团队负责优化资产生命周期，关注物理和财务管理；运维管理员常常使用 CMDB 管理资产配置项及其关系；安全管理员则负责网络资产的监控和安全问题的修复。

谁在做资产管理？

ITAM = Information Technology Asset Management（IT资产管理）
CMDB = Configuration Management Data Base（配置管理数据库）
CSAM = Cybersecurity Asset Management（网络安全资产管理）

	IT管理员	运维管理员	安全管理员
术语	ITAM	CMDB	CSAM
定义	优化和管理IT资产**全部生命周期**：规划、采购、部署、回收、升级、处置等。	管理企业IT环境中的配置项（CI）**相互关系及变更情况**，支持IT服务管理（ITSM）流程，如事件管理、问题管理、变更管理等	**持续发现、盘点、监控网络资产**，时刻积极修复其网络安全问题。
资产/配置项	所有硬件、软件	物理和逻辑资产、系统、应用程序、服务、文档、人员等	任何联网设备、网络服务、安全策略等
目标	资产有效利用、降低成本、提高效率，并确保合规性	保障业务系统无故障长期稳定运行	确保企业网络系统安全稳定运行
关注点	**资产的物理管理、财务管理和合规性**	**资产配置状态及相互关系、变更情况**	**网络空间资产的安全性**

图 2 资产管理的差异

2. 企业资产管理与安全

在信息安全领域有一句名言：你无法保护你看不见的东西。不难理解，资产管理是企业安全的基础，安全管理员在安全事件定位、责任人关联、漏洞影响范围确定、合规性检查等工作中都离不开资产清单，但是，目前许多企业在资产管理方面存在显著不足，因资产管理不善而引发的安全问题也日益严重。当前，企业资产管理通常面临以下几个问题。

第一个是影子 IT 资产问题。影子 IT 资产是指企业员工在未获得 IT 部门批准的情况下使用的个人设备、应用程序和服务，员工这么做通常是为了提高工作的效率和便捷性。随着云服务和移动技术的发展、远程办公以及 BYOD（Bring Your Own Device）政策的普及下，这种现象越发普遍。根据思科的调查结果，大约 83% 的受调查人员承认在企业云环境中使用未经批准的应用程序 [5]。由于这些影子 IT 未被纳入管理，所以存在极大的安全隐患。根据 Gartner 的数据，早在 2020 年就有约 30% 的网络攻击与影子 IT 有关 [2]。所以，企业需要提供能发现和识别影子 IT 的有效手段。然而，如图 3 所示，根据数说安全的调查，58% 的受调查企业在未知资产的发现能力上存在显著不足，只有 11% 的受调查企业认为能够有效识别未知资产 [3]。

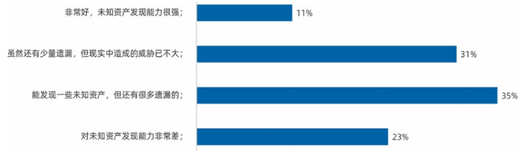


图 3 企业影子资产发现能力调查结果（数说安全）

第二个是云上资产管理问题。随着云计算和云原生技术的发展，云服务因其便捷性和成本效益受到企业的青睐。这使得越来越多的企业将业务迁移至云端，云上资产在企业整体资产中的比重不断增加。据调查，86% 的企业有意向或者已经开始使用云原生服务 [4]。然而，许多企业在管理云上资产方面表现不佳，它们通常仅对云厂商提供的原生管理平台或 API 进行管理，而这对于安全团队来说远远不够，这也促进了另一项技术的发展——云安全态势管理（CSPM）。

第三个是资产数据孤岛问题。如前文所述，一个企业通常会有多个团队参与资产管理工作，这就导致了企业内部可能会存在多个资产管理系统，它们独立分散，没有形成一个统一视图，而这会带来很多部门协作以及使用上的麻烦。比如，在安全团队溯源的时候，需要先从自己的内部网络资产系统获取 IP 信息，再去 IT 部门的资产管理平台查询得到责任人和其他相关信息。而如图 4 所示，根据数说安全 2023 年攻击面管理市场分析报告的调查，18% 的企业没有 IT 资产管理系统，它们可能还采用 Excel 表格这种古老的方式。34% 的企业有多个系统存在 IT 资产管理能力，但是各自为战，数据分散在多个系统中。34% 的企业有统一的 IT 资产管理系统，能有效进行数据整合，只有 15% 的企业的 IT 资产管理系统能同 SIEM/SOC/SOAR 等安全运营系统有效整合，赋能安全运营。

图 4 企业 IT 资产管理能力自我评价调查结果（数说安全）

3. CAASM（网络资产攻击面管理）

如前所述，随着云计算、物联网和远程办公等趋势的发展，组织的网络资产数量和种类迅速增长，且变得更加分散和动态，传统的分散异构的资产管理模式已经难以满足安全团队的需求。在此背景下，网络资产攻击面管理（CAASM）应运而生。CAASM 的核心理念是通过整合现有工具和 API，汇聚多源资产信息，从而获取企业内部和外部所有资产的全面视图，提供统一的资产管理平台，帮助发现安全漏洞并提供修复建议。

CAASM 解决方案可以视作资产管理的延伸，其目标用户不仅限于安全团队，IT 团队及其他部门同样可以从中受益。CAASM 产品通常具备以下功能：全面的资产清单、资产关联图、漏洞管理、安全控制验证、风险优先级排序、风险范围界定、合规性检查以及自动化告警。CAASM 属于攻击面管理（ASM）工具的一种，另两个常见的 ASM 工具是外部攻击面管理（EASM）和数字风险保护服务（DRPS）。图 5 展示了它们之间的区别。

缩写	英文术语	中文翻译	解释	提出时间
ASM	Attack Surface Management	攻击面管理	获取组织**网络资产**，识别漏洞与弱点，制定防御和补救措施。	2019年
DRPS	Digital Risk Protection Services	数字风险保护服务	关注**数字风险**：品牌保护，数据泄露监控，网络钓鱼和欺诈，社交媒体监控	2019年
EASM	External Attack Surface Management	外部攻击面管理	外部攻击者视角，关注企业**暴露**在互联网的资产	2021年
CAASM	Cyber Asset Attack Surface Management	网络资产攻击面管理	内部管理员视角，关注企业内外部**全部**资产	2021年

图 5 ASM 工具区别

4. CAASM 相关厂商和技术

自 Gartner 在 2021 年提出 CAASM（Cyber Asset Attack Surface Management）以来，国内外多家厂商迅速推出了相关的 CAASM 解决方案，以满足企业对资产可视性与安全管理的需求。以下简要介绍两家专门从事 CAASM 的代表性厂商：Axonius 和 JupiterOne。

4.1 Axonius

Axonius 于 2017 年成立，是一家以色列公司，凭借其 CAASM 产品获得了 2019 年 RSA 创新沙盒冠军，并连续四年入选福布斯云 100 强名单。Axonius 并未走大而全的解决方案路线，而是聚焦在资产清点、资产运行及安全状态的监控，它在无须部署任何

代理的情况下，可连接包括网络、端点、身份、漏洞、云管理平台等不同类型共 900 多个数据源。图 6 是其产品的资产页面。

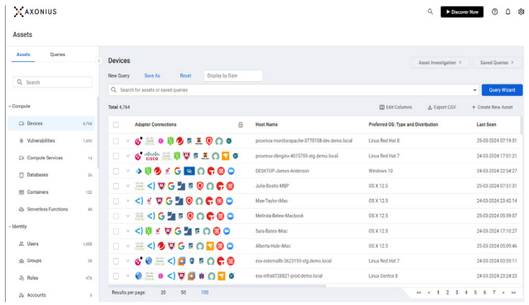


图 6 Axonius CAASM 产品页面

Axonius 将从各类平台获取的资产数据整理为计算、身份、应用、票据、网络、存储、告警与事件七个大类（见表 1），从而提供全面统一的资产清单。Axonius 还通过资产图显示资产关系和依赖关系，Axonius 定义的关系有 Last used by（资产和用户关联）、Has（资产和漏洞关联）、Affected by（资产和其他资产关联）、Accessing（资产与其他资产或用户关联）等。

表 1 Axonius 对资产的分类

分类	包含的资产
计算	设备、漏洞、计算服务、数据库、容器、无服务器函数、计算镜像、配置
身份	用户、组、角色、组织单元、账户、证书、权限
应用	软件、SaaS 应用、应用扩展、用户扩展、应用设置、许可证、费用、活动、业务应用、域名和 URL、应用服务、机密、应用资源
票据	来自 ServiceNow 等应用的票据
网络	网络、负载均衡器、网络服务、网络设备、防火墙规则
存储	对象存储、文件系统、磁盘
警报与事件	警报/事件

Axonius 产品主要有四个功能模块：

（1）软件管理：利用查询语句和事件管理功能，帮助 IT、财务、安全和风险团队自动识别 IT 资产中的所有软件，识别不需要的和未经批准的软件，计算软件支出和成本优化，并降低风险。

（2）云资产合规性：检查云实例是否遵守或偏离行业标准和基准。

（3）漏洞管理：结合必要的上下文来发现影响数字基础设施的漏洞，确定优先级和影响范围。

（4）执行中心：自动化告警和响应，及时分类并修复漏洞和 IT 安全策略问题。

笔者认为，Axonius 的优势在于其强大的多源数据接入能力，能够广泛连接 900 多个不同类型的 IT 与安全系统，帮助企业构建全面的资产视图。此外，Axonius 也有一些特色功能，比如云资产合规性和不依赖第三方平台的本地事件管理功能。

4.2 JupiterOne

JupiterOne 于 2018 年成立，总部位于美国北卡罗来纳州，其 CAASM 产品可以集成 250 多个数据源（见图 7），使用知识图谱对数据进行建模，通过资产、漏洞、业务上下文等关联融合技术，进行攻击面管理分析。

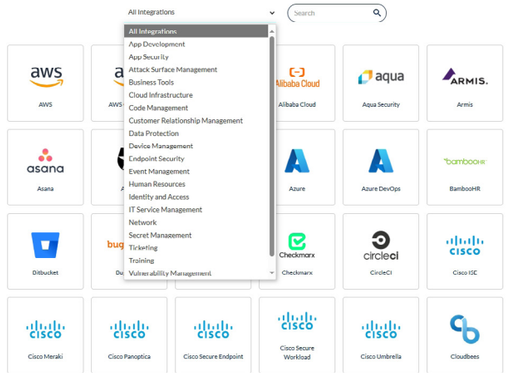


图 7 JupiterOne CAASM 产品集成多种数据源

类似的，JupiterOne 将资产划分为设备、网络、应用、数据、用户五大类，然后定义了 101 个实体，图 8 为实体词云图。可以发现，JupiterOne 把资产分得非常细，除了 Firewall、Host、Endpoint 等比较常规的设备类资产外，像 Key、DataStore、CodeCommit 这类在应用中所定义的一些元素也被当作了资产，并归类为数据资产。JupiterOne 还为这些资产实体定义了超过 26 种关系，表 2 列出了一些关系。实体间通过关系建立起联系，比如 Host -- HAS -> Vulnerability、Gateway -- CONNECTS -> Network、User -- ASSIGNED -> Application。JupiterOne 通过项目 Starbase 从各类服务和系统中（包括云基础设施、SaaS 应用程序、安全控制等）收集以上资产和关系，构建知识图谱，并整合到 Neo4j 数据库中。

图 8 JupiterOne 的资产实体词云图

表 2 JupiterOne 的实体关系

Relationship			
HAS	MANAGES	USES	IDENTIFIED
CONTAINS	EVALUATES	CONNECTS	PERFORMED
IS	MONITORS	TRIGGERS	COMPLETED
OWNS	PROTECTS	EXTENDS	PROVIDES
EXPLOITS	TRUSTS	IMPLEMENTS	CONTRIBUTES TO
IMPACTS	ASSIGNED	MITIGATES	OPENED
MANAGES	DEPLOYED TO		

JupiterOne 产品的主要功能：

(1) 资产清单：集成 200 多个数据源，包括 IT、DevOps、Security、HR 等系统，管理 5000 多个资产类别，包括设备、用户、应用、网络、代码仓库、策略等。

(2) 资产关联：使用知识图谱技术，对资产进行实体建模和关系映射并可视化资产之间的所有联系，还可揭示资产的毒性组合或者会增加风险的隐藏关系。

(3) 资产查询：支持查询语句、全文搜索以及自然语言询问。

(4) 报告生成：使用内置和自定义合规性框架进行自动化审计并收集证据，支持 PCI、CIS、SOC2、NIST、HIPAA 等标准。

(5) 工作流自动化：通过内置和自定义警报，主动通知团队。

笔者认为，JupiterOne 的优势在于对资产更加精细的划分，并采用知识图谱对资产进行建模，从而能对资产进行深入分析和关联，揭示潜在的风险组合及隐藏的风险关系，支持安全团队做出更精准的决策。

5. 小结

企业资产管理是安全团队众多工作的基础，而 CAASM（Cybersecurity Asset Attack Surface Management）的出现有效解决了现有资产管理中的一些弊端。通过集成 API 和现有工具，CAASM 能够提供企业全面的内外部资产视图，并进行弱点分析、响应和修复。目前，CAASM 处于 Gartner 安全运营技术成熟度曲线的期望膨胀期。尽管市场对该技术抱有极高的期望，但其仍处于早期阶段，尚未完全成熟。

现阶段，许多 CAASM（企业资产攻击面管理）解决方案主要集中于整合多个数据源，尤其是来自通用平台的数据源，并对这些数据进行清洗和归一化处理。随后，它们通过基于知识图谱的资产关联技术、基于版本匹配的漏洞验证技术以及基于规则的报警等方式进行风险发现，所采用的技术手段相对传统。未来的 CAASM 应当支持企业内部系统，允许对数据格式进行定制化集成，以更好适应不同企业的特定需求。此外，要逐步引入人工智能和机器学习技术，自动化地识别威胁、优先排序漏洞并进行行为分析。同时，利用更复杂的算法进行细微的威胁预测，将有助于提升企业的安全防护能力。

参考文献

[1]https://csrc.nist.gov/glossary/term/asset.

[2]P. Mukherjee, Short-Guide-On-Attack-Surface-Reduction-v1-2. [Online]. Available: https://www.firecompass.com/wp-content/uploads/2020/12/Short-Guide-On-Attack-Surface Reduction-v1-2.pdf.

[3] 数说安全攻击面管理产品市场分析报告 .https://www.csreviews.cn/wp-content/uploads/2023/05/SSAQgjmgl.pdf.

[4]Scrut Automation, Key Attack Surface Challenges Cloud-Native Companies Are Facing Today. [Online]. Available: https://www.scrut.io/ebooks/key-attack-surface-challenges-cloud-native companies-are-facing-today (accessed: Jun. 25 2023).

[5]The Shadow IT Dilemma - Cisco Blogs(https://blogs.cisco.com/cloud/the-shadow-it-dilemma).

[6]Cyber Asset Attack Surface Management (CAASM) Meaning And Best Practices (tikaj.com).

隐私与安全的深度融合：构建未来数字信任的核心

绿盟科技 创新研究院 顾奇

摘要：在 2024 年的 RSA 大会上，来自比利时普华永道的 Kim Wuyts 分享了题为《隐私融入，让隐私设计与安全设计共生》的议题。她详细讨论了隐私和安全在设计层面的交集，强调隐私不仅仅是数据机密性的保障，更是对业务流程、用户体验、风险管理等方面的全方位考虑。本文将梳理该演讲的主要内容，探讨隐私与安全相互强化的方式及其实践方法。

关键词：隐私设计 威胁建模 隐私增强

1. 背景

随着全球范围内数据隐私问题的日益凸显，隐私保护已从单纯的合规要求，演变为企业系统设计中必须优先考虑的核心要素之一。越来越多的法规（如 GDPR、CPRA）要求企业在数据处理过程中将隐私设计纳入系统开发生命周期的早期阶段。

Kim Wuyts（后简称 Wuyts）认为，尽管如今隐私和安全已经被视为两个独立的领域，但它们在现代系统中的交集依然非常紧密；同时，现有的数字系统通常已对数据安全进行了考虑，因而一个值得思考的方向是如何通过将隐私设计引入安全设计的过程中来提升整体系统的安全性和隐私保护能力。

2. 隐私设计

2.1 隐私的核心要素

隐私设计并不仅仅是确保数据的机密性，如图 1 所示，如 2015 年文献 Protection *Goals for Privacy Engineering* 与 NIST 在 2017 年 [1] 所总结的，隐私的核心要素包含：

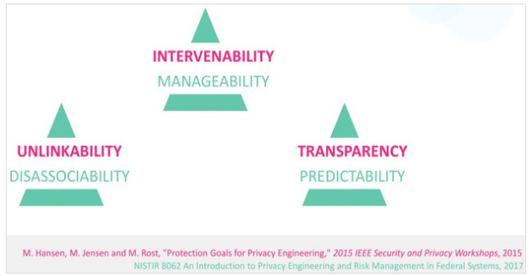


图 1 隐私的三大核心要素

不可关联性 / 可分离性（Unlinkability/ Disassociability）：确保个人及其活动或数据无法轻易被关联到其身份或其他活动。

可干预性 / 可管理性（Intervenability/ Manageability）：用户应能够对其个人数据进行管理和干预，比如能够撤销同意、修改或删除不准确的数据。

透明性 / 可预测性（Transparency/ Predictability）：用户应当能够充分理解其个人数据的使用方式，系统需要向用户清晰地解释数据处理的流程和潜在的影响。

这些要素共同构成了隐私保护的核心，它不仅仅是防止数据泄露，还需要确保用户对其数据的知情权和控制权等。

2.2 隐私与安全

如表 1 所示，尽管隐私和安全的关注点有所不同，但它们可以通过设计实现相互强化。

表 1 安全性与隐私性的主要思考维度

领域	安全性（Security）	隐私性（Privacy）
资产（Assets）	数据的种类、存储技术	更细粒度的控制、数据用途
控制（Controls）	协议、加密、认证	访问控制、用户同意流转、隐私增强技术
行为者（Actors）	用户、角色、攻击者	个人、外部人员、组织本身

安全如何增强隐私：安全的核心是通过技术手段保护数据的机密性、完整性和可用性。这些安全机制（如加密、身份验证等）能够为隐私保护提供基础保障。尤其是机密性，作为隐私的核心要求之一，安全设计中的加密措施可以有效防止未经授权的个人访问敏感数据。

隐私如何增强安全：隐私设计的重点在于减少数据收集和存储，确保数据按预期使用。通过减少不必要的数据收集，隐私设计能够降低数据泄露的潜在风险。比如，数据最小化原则不仅符合隐私保护的要求，也间接减少了系统中可能被攻击者利用的敏感数据，增强了系统的整体安全性。

2.3 威胁建模分析

为了更好地分析和理解系统中的隐私与安全威胁，如表 2 所示，列举了多个知名的威胁模型框架。这些框架帮助企业系统化地识别潜在威胁，并确定适当的应对措施。

表 2 知名威胁建模框架

框架名称	安全性	隐私性
STRIDE	√	
LINDDUN		√
EOP	√	
INCLUDES NO DIRT	√	√
PLOT4AI	√	√
TRIM	√	
STRIPED	√	√
CTM	√	
MITRE PANOPTIC		√

LINDDUN[2] 是最知名的隐私建模框架之一，其具体包含：

关联（Linking）：个人数据可以通过组合其他数据来提取更多的信息；当不同来源的数据组合在一起时，可能会揭示用户的身份或其他隐私信息。

识别（Identifying）：数据可以被绑定到某个自然人，个人可以通过这些数据被直接识别出来，侵犯隐私。

不可否认性（Non Repudiation）：无法否认某个行为或事件的参与，一旦数据被记录，个人很难否认自己与某些活动或事件有关联；这在安全维度上是必要的，但在隐私维度中某些情况下可能是不合理的。

检测（Detecting）：基于观察，可能会推测出额外信息；即使没有明确的个人信息，也可能通过一些推测手段得出个人的敏感信息。

数据披露（Data Disclosure）：个人数据被处理得超出预期的范围，导致敏感信息被泄露，超出了用户的控制范围。

不知情（Unawareness）：用户可能对数据的收集、使用和共享情况不知情，无法有效保护自己的隐私。

不合规（Non Compliance）：系统或组织缺乏或未实施隐私保护的最佳时间，导致不符合隐私法律法规的行为。

除了 LINDDUN，Wuyts 还介绍了其他威胁模型框架，如 STRIDE（专注于安全的威胁建模）。她指出，企业在进行威胁建模时，不能仅仅局限于单一的安全或隐私视角，而应该将两者结合，进行联合分析。通过这种方式，企业可以更全面地识别潜在威胁，并制定更有效的应对策略。

3. 隐私增强方案

面对隐私和安全威胁，Wuyts 强调了隐私增强技术（Privacy Enhancing Technologies, PETs）的重要性。PETs 为隐私保护提供了一系列有效的技术工具，帮助企业在系统设计中主动实施隐私保护。

3.1 常见的隐私设计策略

Wuyts 介绍了几种常见的隐私设计策略，这些策略可以帮助企业在设计阶段就有效地保护用户隐私：

最小化（Minimize）：尽量减少数据的收集和存储，尤其是敏感个人信息。

抽象化（Abstract）：通过模糊数据或聚合数据来替代精确的个人数据，减少隐私风险。

分离（Separate）：将不同类型的数据分开存储和处理，避免集中存储带来的风险。

隐藏（Hide）：通过加密或匿名化技术隐藏个人数据，使其在未经授权的情况下不会被读取。

控制（Control）：提供用户对数据的控制权，允许其撤回、修改或删除数据。

告知（Inform）：确保用户对其数据的处理有充分的知情权和理解。

3.2 PETs 的技术实现

虽然原报告中未对具体技术做展开介绍，但笔者认为下述技术值得关注：

匿名化和去标识化：通过遮掩或移除个人标识信息，确保数据无法直接关联到个人，典型技术包含匿名、多样性、接近性等。

差分隐私（Differential Privacy）：通过引入合适程度的噪声，在保障群体分析有效性的同时，确保个体信息从数据集中被推断出来。

合成数据（Synthetic Data）：一种基于真实数据生成的虚拟数据集，它保持了与真实数据类似的统计特性，但不包含真实的个人信息。

隐私计算（Privacy-Preserving Computation）：隐私计算技术通过确保数据在计算过程中保持加密状态或不暴露原始数据，来实现数据共享和协同计算中的隐私保护。

3.3 PETs 实施的现实挑战

尽管 PETs 为隐私保护提供了强有力的技术支持，但在实施过程中仍面临一些挑战：

技术复杂性：某些 PETs（如差分隐私和零知识证明）在实现上可能非常复杂，企业需要投入大量资源。

性能影响：某些隐私增强技术会对系统性能产生影响，因此在设计时需要平衡隐私保护和系统效率。

用户体验：某些隐私保护措施可能影响用户体验，因此如何在保护隐私的同时保持用户的使用便捷性尤为重要。

企业应在系统开发的早期阶段引入隐私保护措施，避免在后期进行昂贵的修复。

3.4 企业推进方案

Wuyts 在演讲的最后提出，未来隐私和安全将更加紧密地融合在一起。她呼吁企业在设计阶段就将隐私和安全的需求结合起来，而不是将两者视为独立的、事后补救的措施。

为了更好地推动隐私保护和安全的融合，Wuyts 提出了一些分阶段的实施建议：

短期目标（Next Week）：从小处着手，审查现有的数据收集流程，减少不必要的数据收集，并开始评估隐私风险。

中期目标（In 3 Months）：评估团队的隐私技能是否到位，并开展相关的隐私培训，培养隐私设计与安全设计的不同思维模式。

长期目标（In 6-12 Months）：在未来 6 到 12 个月内，企业应全面评估其系统中的隐私和安全策略，并在开发流程中引入隐私工程实践，确保隐私保护从系统设计开始就得到充分考虑。

隐私保护不仅仅是技术问题，还需要在企业内部建立一种隐私文化。隐私保护应成为每个员工的责任，而不仅仅是合规团队或技术团队的任务。通过这种方式，企业可以更加有效地应对未来的隐私挑战，并提升用户对企业的信任。

4. 总结

本文探讨了隐私与安全设计的相互关系，强调两者并非独立的领域，而是可以通过巧妙的设计实现相互促进。现代企业应着眼于将隐私设计融入安全设计中，并通过隐私增强技术（PETs）和威胁建模等方式，确保系统在开发的每个阶段都具备隐私保护能力。

参考文献

[1] https://nvlpubs.nist.gov/nistpubs/ir/2017/nist.ir.8062.pdf.

[2]https://threat-modeling.com/linddun-threat-modeling.

构建精益安全体系：浅析4个攻击面管理技术

绿盟科技 创新研究院 桑鸿庆

摘要：随着企业数字化转型加速，网络资产激增，包括硬件、云服务、物联网设备等，资产安全管理日益复杂。本文聚焦资产攻击面管理（ASM）技术发展，涵盖外部攻击面管理（EASM）、网络资产管理（CAASM）、数字风险防护（DRPS）及持续威胁暴露管理（CTEM）。这些技术从资产发现到威胁监控，为优化企业安全建设提供了新路径。通过建立全面的攻击面管理体系，企业可在有限预算内精准保护核心资产，提升安全运营效率，平衡风险与投资回报。

关键词：资产管理 攻击面 CAASM

1. 引言

随着企业的数字化转型和云计算、物联网等技术的广泛应用，组织拥有的网络资产数量急剧增长，包括硬件设备、软件服务、云服务、物联网设备等。网络资产安全管理变得越来越困难。大多数组织因资产管理不善而引发网络安全事件。资产管理这个老问题，一直都没有被很好地解决，仿佛是一块无法撬动的顽石，无论经过了多少次尝试改进，都未彻底解决，长期困扰着安全运营团队。本文将梳理近年来的资产攻击面管理相关技术，做一些简要的分析，供大家温故而知新。

2. 网络资产安全管理难点

资产管理是企业安全建设中的一个核心痛点，但这一问题至今未能得到有效解决，背后必然存在着诸多难点，从技术视角归纳了一些难点供大家讨论。

2.1 类型多样

随着企业中设备和应用的迅速增加，网络资产管理的复杂性显著提升。企业所使用的设备不仅限于传统的服务器和办公电脑，还包括移动终端、物联网、安全设备、云服务和虚拟化资产等多种类型，企业的资产范围界定逐步趋向于“云、管、边、端”架构体系的四个层级。尤其在疫情期间，远程办公的普及和对云服务的广泛依赖加速了网络设备的多样化，企业不得不迅速适应这些变化以确保业务的连续性。这种设备和应用的多样性不仅增加了管理的复杂性，也对企业的安全防护提出了更高的要求，增加了潜在的安全风险。

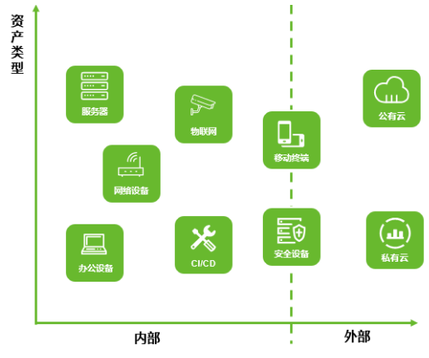


图 1 企业网络资产范围

2.2 数据多源

资产管理面临的另一个主要痛点是数据多源带来的挑战。由于资产信息来源于不同的探针、人工录入和外部系统，往往导致数据的不一致和冲突。例如，相同资产在不同来源中的属性可能存在差异，使管理人员难以获得准确的全景视图。此外，多源数据整合的复杂性使实时更新变得困难，进而影响，资产的准确性和及时性。缺乏标准化的格式和流程也加剧了这一问题，增加了数据清洗和验证的工作量，最终导致决策的延误和管理效率的降低。数据不一致性的问题使融合变得更加困难。

2.3 动态变化

资产的类型、数量、配置和风险都会随着时间而变化，给 IT 和安全团队带来巨大的管理负担。例如，IT 团队不仅要跟踪这些设备的存在，还需要确保它们始终处于最新的安全状态，如安装了最新的补丁和安全更新。随着网络资产的持续增长，产生的“盲区”也使团队难以全面了解和控制所有资产的风险状况。同时，网络威胁的态势不断演变，网络犯罪分子持续寻找新的漏洞利用方法，企业面临的威胁愈加多样和复杂。这种动态变化使得网络资产管理的复杂性进一步加剧，迫使安全团队必须在极短的时间内做出反应。

2.4 自动化

我们总是谈企业安全建设智能化，但是资产安全管理实际情况是自动化数据采集，很多企业都没有完全覆盖，使每个安全运营的“老师傅”都有一套自己的方法论。自动化技术的缺乏使资产管理的各个环节无法实现实时监控和快速响应，进一步加大了企业在面对复杂网络环境和不断演变的威胁时的脆弱性。这个自动化并不是面向一个人的自动，而是在企业资产管理整个流程中的自动化，所以这就面临着复杂性、数据不一致、缺乏标准化流程、技术整合难度大、文化抵触、合规要求和投资成本高等多重挑战。

3. 攻击面相关技术介绍

3.1 攻击面管理介绍

在网络安全的对抗中，攻击者与防守者之间一直都是一个不对等的关系，主要包括资产视角、主动性、资源消耗、复杂性以及时间因素的差异。这种不对等使防守者处于较为被动和困难的局面，而攻击者则相对具备更多优势。于是在安全技术的不断演进中，攻击面管理（Attack Surface Management，ASM）就出现了，其核心思想是从攻击者视角对全类型资产暴露面进行发现与管理的过程，通过对攻击面进行深入分析和有效管理，显著提高企业网络安全运营的响应速度。

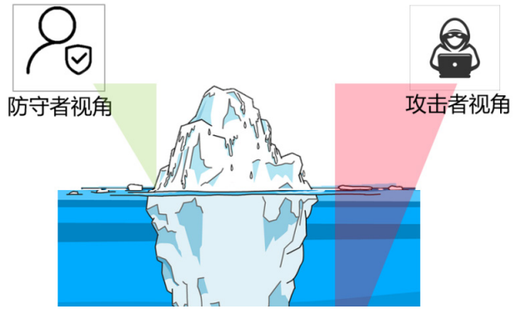


图 2 攻防双方视角的不对等

攻击面管理的发展历程伴随互联网的迅速扩展而不断演进。早期，网络安全工具主要通过简单的搜索引擎查询，如 21 世纪初期的“Google Dorks”工具，以及 Shodan 等网络服务搜索引擎，来识别暴露的数字资产。随着威胁日益复杂，安全研究者开发了更高级的工具，如 Zmap 和 Masscan，用于大规模快速扫描。到 21 世纪 10 年代，ASM 工具逐步从简单的资产发现，演变为能够全面管理攻击面，包括发现、分析、优先级排序、修复和持续监控潜在的安全风险。近年来，随着云计算、物联网（IoT）等新技术的兴起，ASM 工具必须适应新的威胁向量，逐步发展为涵盖外部攻击面管理（EASM）、网络资产攻击面管理（CAASM）、持续威胁暴露管理（CTEM）等更为全面的解决方案，如图 2 所示，Gartner 2024 年的技术成熟度曲线可以看出 CAASM 和 EASM 分别处于高峰期和低谷期。

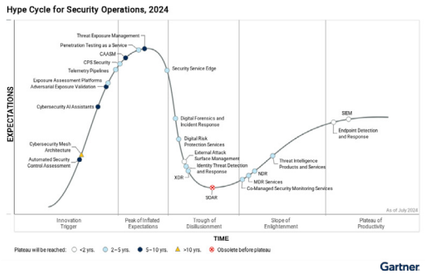


图 3 Gartner 2024 年技术成熟度曲线

3.2 相关技术介绍

外部攻击面管理（EASM）

EASM 专注于识别和管理企业的外部互联网资产及其漏洞。这些资产包括暴露的服务器、公共云服务配置错误以及第三方合作伙伴的软件代码漏洞等。EASM 的主要目的是帮助组织识别外部的攻击面，提供攻击者视角，找出可能被利用的外部风险。

EASM 适合那些主要面临互联网暴露资产和外部攻击风险的企业，能够通过外部资产扫描和漏洞检测发现潜在的威胁。对于需要持续监控外部资产并进行威胁狩猎的安全团队来说，EASM 是必不可少的工具。

网络资产攻击面管理（CAASM）

CAASM 的重点是提高对组织内部和外部资产的可见性。通过与现有的 IT 工具集成，CAASM 能够汇总数据，帮助组织全面了解其资产状态。CAASM 还可以识别安全控制中的漏洞，并通过自动化措施解决问题。与 EASM 不同的是，CAASM 更多地依赖现有的内部数据源，而非扫描外部资产。

CAASM 适合资产管理较为混乱、缺乏全面资产可见性的组织，特别是在跨部门和跨工具的复杂 IT 基础设施环境中。它通过自动化手段汇总资产信息，优化资产管理流程，是想要提升内部和外部资产透明度的企业的理想选择。

数字风险防护服务（DRPS）

DRPS 主要针对品牌保护、第三方风险评估以及外部威胁的发现。这类技术服务可以监控表面网络、社交媒体、暗网和深网，以识别可能针对企业资产的威胁。它还提供对威胁行为者、策略和攻击流程的上下文信息，帮助企业应对外部风险。

DRPS 适用于需要品牌保护、社交媒体威胁检测以及数据泄露防护的企业，特别是在合规性要求高、品牌声誉管理和第三方风险评估方面有需求的组织。它可以帮助企业监控外部网络，及时发现和应对潜在的外部风险。

持续威胁暴露管理（CTEM）

CTEM 不是一个技术，它是方案和架构，更强调通过持续检测、评估和管理暴露面来提高组织对威胁的响应能力。CTEM 涉及整个暴露面管理的周期，从资产可见性到漏洞评估，再到模拟攻击测试，以确保组织能够持续识别和优先处理暴露面上的风险。CTEM 通常需要整合

EASM、CAASM 和 DRPS 等技术工具，以形成全方位的威胁管理方案。

CTEM 适合需要持续威胁暴露管理的组织，它提供从资产发现到威胁模拟的全流程管理能力。如果企业已经具备了攻击面管理的基础能力，CTEM 就能够确保对整个攻击面进行全面、持续的监控和快速响应。

3.3 技术对比

攻击面管理是一个复杂的过程，因此衍生出了 EASM、DRPS、CAASM 和 CTEM 这四种技术来应对不同的需求和场景。如图 4 所示，它们在技术能力上有共同点，比如资产管理和风险优先级排序，但各自侧重不同。EASM 核心外部探测和情报收集，最容易部署，适合快速识别外部风险；DRPS 专注于数字资产风险监控，更加垂直；CAASM 则以网络资产为核心，整合内部资产管理，提供统一的资产风险视图；而 CTEM 是一个全面的解决方案，覆盖全生命周期的威胁管理，需要技术和业务流程结合，实现持续的风险监控和优化。总的来说，这些技术共同构建了一个完整的攻击面管理体系，虽然这些技术各有侧重，但它们的共同目标都是帮助企业有效管理和降低其攻击面风险。企业可以根据自己的业务情况和侧重选择对应的能力建设。

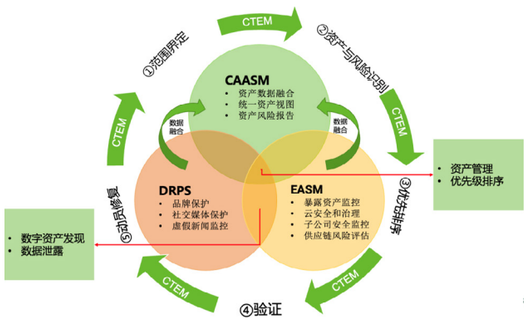


图 4 四种攻击面管理技术的关系

4. 总结

本文分析了近年来资产攻击面管理相关技术的发展，包括 EASM、DRPS、CAASM、CTEM 等相关技术和解决方案。在经济大环境不佳的情况下，企业的安全建设投入往往会受到影响。然而，当前的安全体系是缺失与过载并存，企业可能缺乏对关键资产的保护，同时又在无效或冗余的安全工具上投入过多。这为建立精益安全体系提供了切入点：通过精准识别和优先保护核心资产，优化安全资源的配置，企业可以在减少不必要投入的同时提升整体安全效果。所以如果组织使用 CAASM 建立起资产的整体视图，就可以量体裁衣，能够在有限的预算下，通过资产视图驱动的安全建设策略、有效的资产管理和持续的威胁监控，构建更高效、精益的安全体系，实现风险最小化与投资回报的平衡。

参考文献

[1] https://www.paloaltonetworks.com/cyberpedia/what-is-attack-surface-management.

[2] From Chaos to Control: Simplifying Cybersecurity Asset Management.

[3] https://www.armis.com/success/simplifying-cybersecurity-asset-management-wp-success/.

[4] https://www.cybermindr.com/cyber-security/the-birth-death-of-asm/.

https://socradar.io/gartner-recognized-socradar-as-both-easm-and-drps-vendor/.

可信计算之完整性度量架构

绿盟科技 创新研究院 杨博杰

摘要：随着网络安全威胁的不断升级，可信计算逐渐成为保护信息系统安全的重要技术。可信计算中的完整性度量架构（IMA）是用于验证系统软件和硬件是否被篡改的重要手段。IMA 通过度量系统中所有加载的文件并将度量结果存储到内核中的度量列表中，结合 TPM 硬件技术，可以实现远程证明，保障系统的可信性。本文介绍了 IMA 的基本功能、组件及其与 TPM 的结合使用，分析了其在系统安全中的作用以及存在的局限性。通过对 IMA 与 EVM（扩展验证模块）配合使用的探讨，进一步展示了该技术在提升操作系统完整性检测能力中的重要性。尽管 IMA 在保护文件完整性方面具有一定优势，但仍需结合其他安全策略以应对更复杂的攻击场景。

关键词：完整性度量架构（IMA）可信计算 远程证明

1. 背景

可信计算场景下，证明一个系统是否可信的信息由系统自身产生，验证方需要判断系统哪一部分给出的答案才是值得相信的。返回可信信息的程序本身，以及运行这些程序的 BIOS、内核，都是不可信的，因为它们都可能被篡改。

2004 年，IBM 在 第 13 届 USENIX Security Symposium 上发表论文 *Design and Implementation of a TCG-based Integrity Measurement Architecture,* 第一次提出了 IMA（全称 Integrity Measurement Architecture，完整性度量架构）的概念。IMA 将验证可信的重点放在了软件运行栈的完整性状态，而不是传统的基于数字签名做验证。

IMA 作用下，所有加载进 Linux 系统的可执行文件，在执行前都要被度量，并且度量结果受 TPM 保护。IMA 的加入，成功地将 TCG 的可信度量从 BIOS 扩展到了应用层。

2.IMA 简介

2.1 IMA 的功能

（1）收集：在访问文件前对文件进行度量，计算文件的哈希值。

（2）存储：将度量结果存储到内核维护的度量列表中。

（3）证明：如果存在 TPM，使用 TPM 密钥对 IMA PCR 签名，实现远程证明。

（4）评估：度量文件，并与一个存储在扩展属性中的参考值做比较，控制对本地文件的访问。

（5）保护：保护文件的安全扩展属性，防止离线攻击。

（6）审计：将度量结果写到系统日志中，用于审计。

保护的功能需要用到 EVM（Extended Verification Module，扩展验证模块）。EVM 能够检测到文件安全扩展属性的篡改，防止针对安全扩展属性的离线攻击。EVM 利用一系列安全扩展属性计算 HMAC-sha1 哈希值，存为 “security.evm”，后续在访问文件时利用该属性做验签操作，检测篡改。

2.2 IMA 组件

（1）IMA-measurement：内核完整性子系统的组件之一，提供度量能力。维护完整性度量列表，用于远程证明。

（2）IMA-appraisal：提供评估能力，对本地文件进行完整性校验。对文件进行度量，并与一个存储在安全扩展属性 “security.ima” 中的参考值作比较，如果值不一致则拒绝访问。扩展了 “安全引导” 的概念，在允许系统访问文件前验证文件的完整性。

（3）IMA-audit：提供审计能力。将文件度量值写入系统审计日志中，便于后续的分析。

度量和评估不同：度量只记录当前的状态，不会影响文件的使用。而评估会在发现文件度量值不一样的时候，拦截对文件的访问。

2.3 与 TPM 结合使用

在设备配备 TPM 的情况下，IMA 在计算出文件度量结果后，可以将结果扩展到 TPM 的 IMA PCR 中。扩展是单向的 hash 操作，结合 TPM 的硬件安全性，使得存入 TPM PCR 中的值不可被篡改，保证了度量结果的真实性。此后该结果可以用在远程证明中。

远程证明的目的是将硬件平台的可信传递到网络环境中。为了实现这个目标，TPM 芯片在需要的时候，要能够对外证明当前设备和运行状态的完整性。

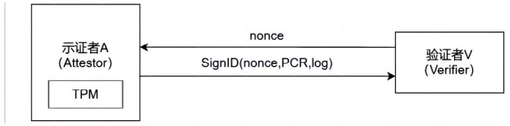


图 1 “挑战 - 应答” 协议

远程证明的协议过程如图1所示。该协议称为“挑战 - 应答”协议。

挑战者（验证者 V）向证明者（示证者 A）发送一个挑战证明的消息以及一个随机数 nonce。该随机数的存在可以防止重放攻击。

证明者把经过身份密钥或签名密钥加签的 PCR 结果、度量日志以及随机数返回给挑战者。挑战者拿到返回后，先用公钥对 PCR 值进行验签，验签通过后，将度量日志的每一项与预期值进行对比，以此判断平台是否可信。

在远程证明过程中，TPM 可以对 IMA PCR 加签，将结果与内核中的度量列表一起发送给对方，以此验证平台的运行完整性。

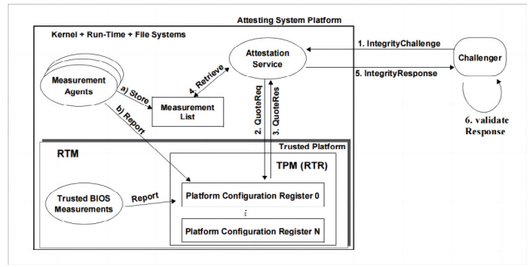


图 2 基于 TPM 的完整性度量

图 2 展示了在 TPM 硬件基础上实现的完整性度量，以及基于该架构实现远程证明的过程。

度量代理（Measurement Agents）对文件做度量，获取度量值，将结果存入内核中的度量列表中，同时将结果存入 TPM 的 PCR 中。

挑战者向平台上的证明服务发起完整性挑战，请求获取度量列表和 TPM 签名的度量列表摘要。

证明服务接收到请求后，从 TPM 中取出 TPM 加签的摘要，以及内核中的度量列表，一起返回给挑战者。

挑战者利用返回的度量结果，验证当前平台的运行完整性，确认是否可信。

3.IMA 的局限性

IMA 依赖在系统调用上加入钩子。只有当指定系统调用（如 execve、open）触发的时候，才会进入 IMA 完整性计算步骤。因此，在文件被加载到系统并运行前，恶意攻击者仍有机会实现篡改。在引入 EVM 后，这部分攻击可以被阻止。

完整性度量功能只对文件的状态做记录，而不做任何拦截。在仅启用该功能的情况下，无法有效地阻止攻击的发生，获取到文件写权限的攻击者完全可以篡改文件内容。只有当度量结果被报告并被检测到异常时，才能对当前系统上的未授权修改做出响应。

评估功能让文件加载前，文件的完整性被校验，能够有效拦截针对文件的修改导致的错误或者恶意代码执行。但是并不能阻止离线攻击（攻击者可以将硬盘取下，到其他设备上修改后，在重新装载）。这就导致有可能发生文件恶意修改，反向利用 IMA 评估阻止某些程序的运行，影响系统正常功能。

为了加强保护，可以结合硬盘加密。在硬盘被加密的情况下，即使攻击者通过物理方式接触到了硬盘，由于无法获取密钥解密硬盘，也就无法对其中的内容做针对性修改。

总体而言，IMA 是针对可信计算完整性检测的一个增强安全能力，但为了确保足够的安全，配合使用其他安全策略是很有必要的。

4. 总结

IMA 作为内核完整性子系统的一部分，强化了操作系统的运行安全。未来，随着安全威胁的不断演变，IMA 将在抵御供应链攻击、保护固件安全、增强勒索软件防护、物联网安全、云安全、可信计算等领域发挥越来越重要的作用。

参考文献

[1]https://sourceforge.net/p/linux-ima/wiki/Home/#features

[2]Sailer, Reiner & Zhang, Xiaolan & Jaeger, Trent & van Doorn, Leendert. (2004). Design and Implementation of a TCG-based Integrity Measurement Architecture. 223-238.

IBM Security 《2024年数据泄露成本报告》解读

绿盟科技 创新研究院 刘文新

摘要：本文是对《IBM 2024 年数据泄露成本报告》的解读。报告显示，全球数据泄露平均成本持续上升，达到 488 万美元，较 2023 年增长约 10%。医疗行业虽成本下降，但仍居行业之首。恶意内部人员和凭证盗用是主要攻击媒介，而云环境成为高风险地带。影子数据泄露导致更高成本和更长响应时间。报告强调，企业应将安全融入 DevOps 流程，应用 AI 和自动化技术，加强员工培训，并与执法部门合作，以降低风险和成本。

关键词：数据泄露 数据泄露成本 IBM

1. 概述

2024 年 7 月，IBM 发布了《2024 年数据泄露成本报告》（*Costofa Data Breach Report* 2024，以下简称报告）[1]。该报告调研了 2023 年 3 月至 2024 年 2 月期间，16 个国家和地区，17 个不同行业的 604 家受数据泄露影响的企业，分析、研究了数据泄露事件发生的原因及后果，探讨了预防此类事件发生的技术和措施。今年已是该报告连续发布的第 19 年。

报告的全部内容为对 14 个不同主题的分析，包含“全球关注重点”“初始攻击媒介和根本原因”“数据泄露生命周期”“识别泄露事件”“安全 AI 和自动化”“发生泄露后提高价格”“业务中断”“恢复时间”“增大或减少泄露成本的因素”“勒索攻击的成本”“报告泄露事件和监管罚款”“数据安全”“大规模泄露”“安全性投资”。

本文将从以下三个方面对报告进行解读。首先，笔者会将报告中提到的 14 个不同主题的内容进行整合、归纳，总结为 7 个主要观点，并针对各观点进行分析。其次，笔者结合报告内容给出帮助企业降低数据泄露风险及成本的最佳实践措施。最后，对报告及该解读文章进行总结。

相关名词解释：

数据泄露平均成本：企业在遭受数据泄露事件后产生的直接和间接成本。其中，直接成本包含聘请安全团队，购买安全服务、产品，罚单等；间接成本包含因事件影响的营业额、客户损失的价值等。

初始攻击媒介：攻击者在试图非法访问 IT 系统和敏感信息时使用的方法，一般通过利用网络、系统或应用程序中的漏洞来实现 [2]。

数据泄露事件响应时间：企业识别和遏制数据泄露事件所花费的平均调查时间（MTTI，Mean Time To Investigate）与平均遏制时间（MTTC，Mean Time To Contain）之和。平均调查时间指从确认一个安全事件到开始调查其原因和解决方案的平均时间；平均遏制时间指安全团队找到威胁者并阻止其进一步进行恶意利用的时间 [3]。

云环境：本文提到的云环境泛指公有云环境、私有云环境和跨云环境。

影子数据：影子数据指的是那些未经正式管理渠道且可能对信息技术部门而言处于隐匿状态的数据。这种情况通常源于员工使用未经许可的应用程序进行数据共享，或是私自将数据上传至非官方云存储解决方案中。

DevOps：DevOps 是一套实践、工具和文化理念，可以实现软件开发团队和 IT 团队之间的流程自动化和集成。它强调团队赋能、跨团队沟通和协作以及技术自动化 [4]。

2. 观点分析

本节将报告的 14 个主题提炼为 7 个核心观点，并对每个观点进行分析，全部主题内容请参考完整报告。

观点 1：2024 年数据泄露平均成本创下历史新高，达到 488 万美元，较 2023 年增加约 10%；医疗行业数据泄露平均成本达 977 万美元，较 2023 年降低约 11%，但仍是数据泄露平均成本最高的行业。

图 1 展示了数据泄露平均成本的增长情况。业务损失、应对措施成本的增加，不断演进的攻击手段和技术，以及日益严格的法规和监管要求，共同推动了数据泄露成本的增加。2024 年数据泄露平均成本上升至 488 万美元，比 2023 年的 445 万美元增加 13 万美元，达到历史新高。488 万美元这一数字凸显了数据泄露事件的严重性，以及对企业、对国家乃至对全球经济产生的重大影响。

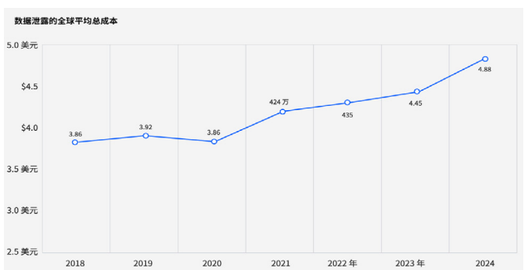


图 1 数据泄露平均成本（单位：百万美元）

图 2 展示了不同行业数据泄露平均成本情况。2024 年医疗行业数据泄露平均成本下降约 11%，降至 977 万美元，但仍是全行业最高。由于医疗行业数据价值高、敏感性强，数据存储环境复杂多样，合规性、监管压力较大等，导致医疗行业已连续 14 年成为数据泄露平均成本最高的行业，医疗行业在数据保护和防范泄露方面面临着前所未有的风险和高昂的成本。

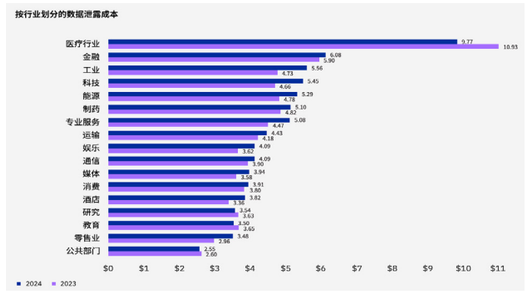


图 2 不同行业数据泄露平均成本（单位：百万美元）

虽然国内暂无针对数据泄露事件平均成本进行统计的相关报告，但威胁猎人发布的《2024 年上半年数据泄露风险态势报告》[5] 中指出，上半年国内数据泄露事件数量 Top5 行业分别为银行、电商、消费金融、保险、快递。其中，银行行业数据泄露事件数量高达 2961 起，成为数据泄露事件数最多的行业。由于事件数据是影响数据泄露成本的重要因素，笔者合理猜测，与国外数据泄露成本最高行业为医疗行业不同的是，国内数据泄露成本最高的行业是银行。银行的数据规模大、数据应用范围广、生态应用和创新多，导致银行成为数据信息敏感程度最高的行业之一，因此也成为数据泄露事件发生量最高的行业 [6]。

国内外数据泄露成本最高行业之间存在差异，背后反映了行业性质、法规要求和数据敏感性等方面的差异。银行业是国内经济中枢，持有大量客户的金融信息，一旦泄露会威胁到经济安全和个人财产。而在美国，医疗数据如病历、健康记录等被认为是极其敏感的隐私信息，泄露会引发身份盗窃、保险欺诈等问题。美国《健康保险可携性和责任法案》（HIPAA）等法规对医疗数据保护有严格的规定，违规泄露会带来高额罚款和法律责任，医疗数据的复杂性也加剧了处理成本。因此，国外医疗行业的数据泄露成本最高。

观点 2：造成数据泄露平均成本最高的初始攻击媒介是恶意内部人员；数据泄露事件频率最高的初始攻击媒介是凭证盗用或泄露；需要响应时间最长的初始攻击媒介也是凭证盗用或泄露。

报告中提到的初始攻击媒介指引起数据泄露事件发生的主要原因或攻击手段，共包含 11 类：商业电子邮件泄露、网络钓鱼、恶意内部人员、凭据盗用或泄露、社会工程、未知 0 Day 漏洞、已知但未修补漏洞、意外数据丢失和设备丢失或被盗、物理安全损害、云配置错误、系统错误。

如图 3 所示，凭证盗用或泄露成为引发数据泄露事件最常见的初始攻击媒介，约占 16%；紧随其后的是网络钓鱼，约占 15%；云配置错误排第三，约 12%。凭据盗用或泄露成为引发数据泄露事件最常见的原因，归因于以下几点：首先，很多企业和用户存在使用默认口令和弱口令的情况，增加了被攻击者暴力破解的风险。其次，企业和用户对凭证的管理和保护不足，存在如凭证硬编码、未启用多因子认证等不安全行为，都会给攻击者提供可乘之机。

恶意内部人员成为引起数据泄露平均成本最高的初始攻击媒介，高达 499 万美元；网络钓鱼和商业电子邮件泄露并列第二，达 488 万美元；凭据盗用或泄露排名第三，达 481 万美元。恶意内部人员成为数据泄露成本最高的初始攻击媒介，揭示了内部安全管理的重大挑战，凸显了内部威胁的隐蔽性和破坏性，还强调了加强员工安全意识、背景调查及访问控制的重要性。

值得注意的是，由于部分企业在补丁管理和漏洞修复方面存在滞后性，未能及时修补已公开的安全威胁，存在 Nday 漏洞未被修复的情况。如图 4 所示，仍有超 6% 以上的数据泄露事件源于尚未修补的 Nday 漏洞，此类数据泄露事件的数据泄露平均成本达 433 万美元，排名第六。此类初始攻击媒介利用成本较低、风险较大，因此企业应提高对此类攻击媒介的关注和重视。

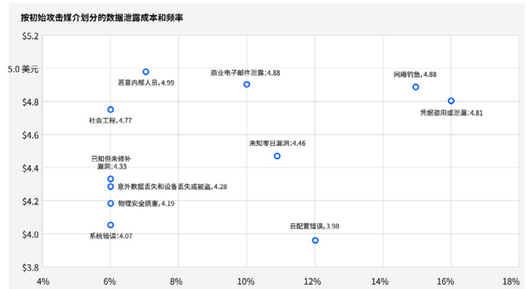


图 3 不同初始攻击媒介下的数据泄露成本和频率（单位：百万美元）

图 4 展示了发生数据泄露事件后企业响应不同初始攻击媒介导致的安全事件所需的时间。其中，企业响应由凭证盗用或泄露引发的数据泄露事件耗费时间最长，约 292 天；恶意内部人员与网络钓鱼分别位居第二、第三，响应此类时间所耗费的平均时长分别为 287 天及 261 天。

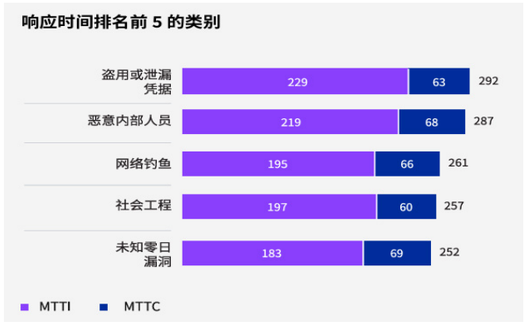


图 4 不同初始攻击媒介对响应数据泄露事件时间的影响（单位：天）

由凭证盗用或泄露引发的数据泄露事件通常隐蔽性较强，攻击者可能长期静默地利用这些凭证进行非法活动，而企业往往难以在短时间内察觉。同时，此类事件检测难度较大。由于凭证泄露较为隐蔽，企业需要通过复杂的监控和检测手段才能发现潜在的异常活动。而且，此类事件影响范围较广，一旦凭证泄露，攻击者可能利用这些凭证访问多个系统或应用，导致数据泄露的影响范围扩大。

观点 3：企业自主披露数据泄露事件的频率在增加，但仍有约四分之一的数据泄露事件是由攻击者披露的。而这类数据泄露事件的平均数据泄露成本最高，响应时间最久。

如图 5 所示，约 42% 的企业安全团队能够自主识别、披露正在经历的数据泄露事件，这一数字比 2023 年报告中的 33% 有所提高，但仍有约 24% 的数据泄露事件是由攻击者披露的。企业正逐步提升自主识别与披露数据泄露事件的能力，体现了对数据安全重视程度的增强及透明化管理的趋势。

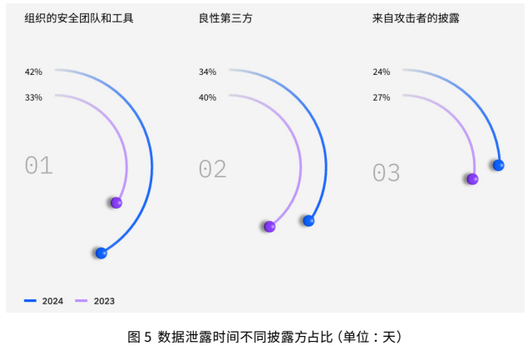


图 5 数据泄露时间不同披露方占比（单位：天）

图 6、图 7 分别展示了数据泄露事件披露方与数据泄露平均成本、数据泄露时间耗时之间的关系。能够发现，由攻击者披露的数据泄露事件的平均数据泄露成本最高，高达 553 万美元，响应时间最久，长达 289 天。

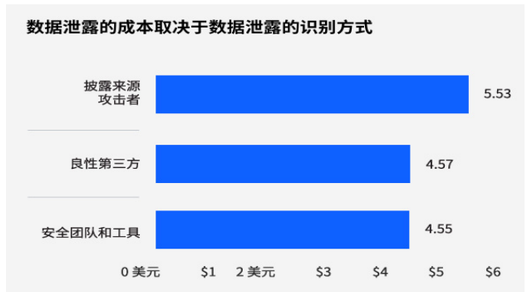


图 6 数据泄露披露方与平均成本关系（单位：百万美元）

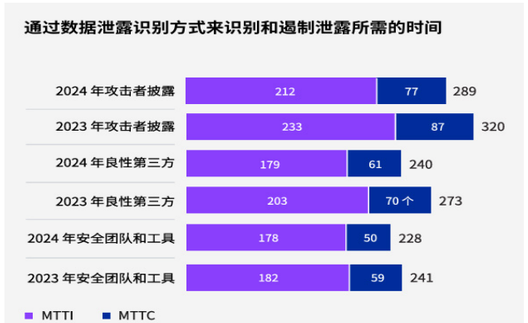


图 7 数据泄露披露方与响应时间关系（单位：天）

由攻击者披露的数据泄露事件之所以泄露平均成本高、响应时间久，主要有以下几个原因。首先，攻击者往往针对高价值数据发起攻击，这些数据可能涉及客户隐私、商业秘密或敏感财务信息，一旦泄露，将对企业造成重大经济损失和声誉损害。其次，攻击者在披露事件前可能已经长期、静默地对企业数据进行窃取且已达到某种目的，导致企业难以察觉与应对，增加了事件响应的难度。最后，这类数据泄露事件往往伴随广泛的媒体报道和公众关注，这也可能导致决策过程复杂化，延长了响应时间。

观点 4：企业采用 AI 和其他自动化安全技术能够有效降低平均数据泄露成本，缩短数据泄露事件响应时间。

报告指出，约有 67% 的企业均有不同程度地使用 AI 和自动化提高企业的安全能力建设，与 2023 年相比，增幅约为 9.1%。

如图 8 所示，企业使用 AI 和自动化安全技术程度越高，其数据泄露平均成本就会越低。从图 8 中可以看出，若企业未曾使用 AI 和自动化安全技术，则其面临数据泄露事件时的平均成本达到 572 万美元，比广泛使用 AI 和自动化安全技术的企业高 188 万美元。另外，广泛使用 AI 和自动化安全技术的企业在响应数据泄露事件时所花费的时间比完全不使用的企业平均快了近 100 天，参考图 9。

企业应用 AI 和自动化安全技术首先能够提高对数据泄露事件的预警能力。AI 和自动化安全技术能够高效地识别常规、传统的威胁和异常行为，更能够精准识别潜在的、复杂的攻击行为，实现快速、早期预警。此外，AI 和自动化安全技术能够加快企业对数据泄露事件的响应。基于 AI 和自动的安全响应机制一旦被触发，便能够立刻作出响应，实现系统隔离、攻击路径阻断等，实现对数据泄露事件的遏制，降低数据泄露事件带来的成本。

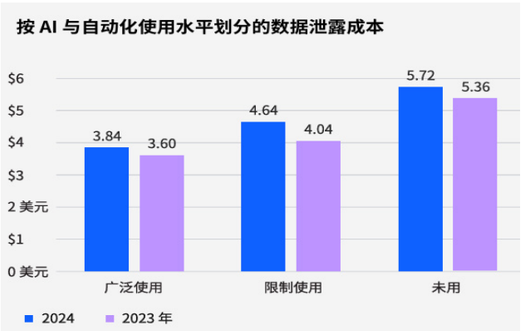


图 8 AI 与自动化使用程度与数据泄露成本的关系（单位：百万美元）

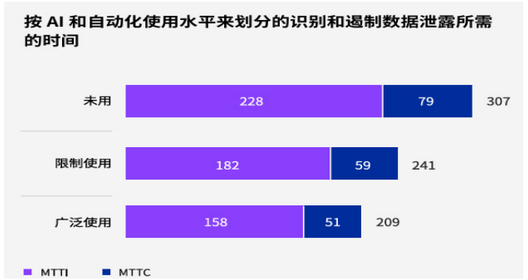


图 9 AI 与自动化使用程度与时间响应时间的关系（单位：天）

报告还对影响数据泄露平均成本的因素进行了分析。如图 10 所示，能够降低数据泄露平均成本的三个主要因素分别是员工培训、AI 驱动的洞察及安全信息和事件管理。如图 11 所示，安全系统复杂性、安全技能短缺和第三方违规是增加数据泄露成本最主要的三个因素。

数据泄露成本的高低受多重因素影响。员工培训能够增强员工的安全意识，减少因人为疏忽导致的数据泄露事件。而 AI 的驱动能及时发现潜在的安全威胁，并通过自动、智能化的方式迅速响应，防止事态扩大，从而控制成本。最后，安全信息和事件管理（SIEM），能够整合并分析安全日志，实现安全事件的快速定位与处置，同样有助于降低数据泄露事件的成本。

相反地，安全系统的复杂导致运维难度增加，可能会降低响应效率，间接推高泄露成本。安全技能短缺意味着在应对安全事件时可能缺乏足够的专业知识和资源，难以有效遏制事态，增加成本。而第三方泄露，特别是供应链中的泄露，由于其隐蔽性和难以控制性，往往造成更广泛的影响，显著增加数据泄露的修复和赔偿成本。

图 10 降低平均数据泄露成本的因素（单位：百万美元）

图 11 增大平均数据泄露成本的因素（单位：百万美元）

观点 5：执法部门的介入能够有效降低勒索事件的平均成本和事件的响应时间。

如图 12、图 13 所示，63% 的勒索软件受害企业选择让执法部门介入，但仍有 37% 的勒索软件受害企业不会选择让执法部门介入。当勒索事件发生时，不让执法部门介入的企业，勒索事件的平均成本达 537 万美元。而让执法部门介入的企业，勒索事件的平均成本为 438 万美元，相差近 100 万美元。当企业遭受勒索攻击时，让执法部门介入是一个很好的选择，能够帮助企业降低约 18% 的平均成本。

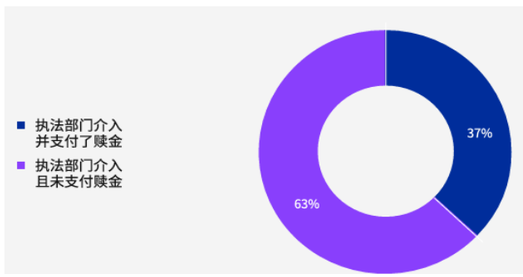


图 12 涉及执法部门但支付了赎金的勒索软件受害者占比

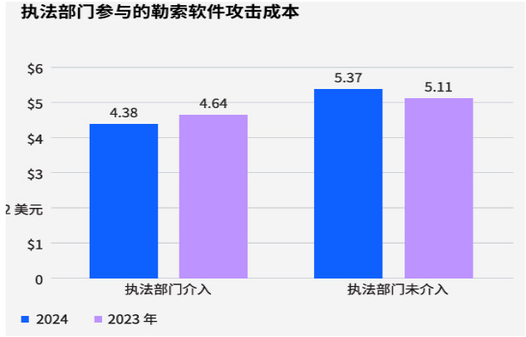


图 13 执法部门参与的勒索软件攻击成本（单位：百万美元）

图 14 展示了执法部门的介入对企业响应勒索事件所花费时间的影响，在执法部门的介入下，识别和遏制勒索软件泄露的总时间缩短 16 天，总时间为 281 天，而执法部门未介入时间为 297 天。

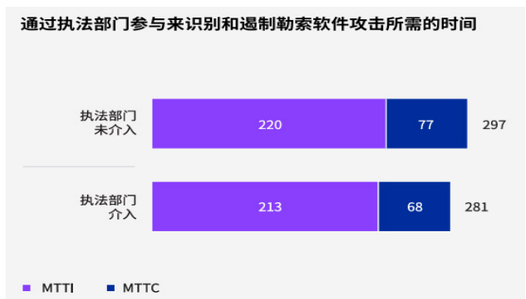


图 14 执法部门的介入对企业响应勒索事件时间的影响（单位：百万美元）

执法部门的介入在应对勒索事件中扮演着至关重要的角色。其权威性和专业性能够有效震慑犯罪分子，从而降低企业所需承担

的数据泄露平均成本。同时，执法部门具备更加高效的调查能力和资源，能够显著缩短事件的响应时间，减少受害方因长时间的僵持而要面临的进一步损失。此外，执法部门还能够为受害方提供法律援助，帮助受害方通过合法途径恢复权益，维护社会秩序与安全。

观点 6：云环境成为数据泄露事件的高发地带，云环境中数据泄露事件的平均成本与响应时间远高于本地环境。

如图 15 所示，数据泄露事件中被泄露的数据约 80% 存储在与云相关的环境中。其中，40% 是跨云环境，25% 是公有云环境，还有 15% 存储在私有云环境中。企业的云化使得企业网络边界逐渐模糊，传统的边界型安全机制、设备无法有效对云环境的安全进行加固。此外，云环境中的安全配置十分烦琐，容易出现简单配置错误导致的安全问题。因此，云环境比本地环境更易发生数据泄露风险，而云环境数据泄露事件的平均成本也会更高、响应时间需要更久。

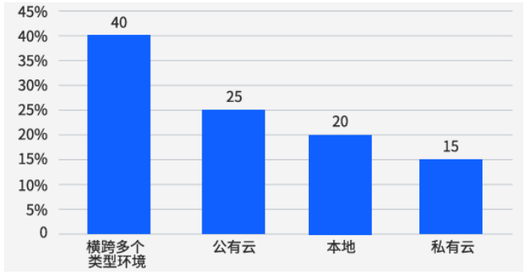


图 15 不同环境下数据泄露事件发生的频率

如图 16 所示，不同环境下数据泄露平均成本较 2023 年均有不同程度的增加。其中，公有云环境下的数据泄露成本最高，约 517 万美元；第二是跨云环境，约 503 万美元；第三是私有云环境，约 433 万美元。本地环境数据泄露成本最低，约 418 万美元。图 17 展示了不同环境下数据泄露事件响应时间的情况。其中，混合云环境发生数据泄露事件的平均识别和响应时间最长，高达 283 天；公有云环境及私有云环境紧随其后，分别为 268 天和 247 天。

云环境正日益成为网络攻击的焦点，混合云环境下的安全和数据泄露防护更是迫在眉睫。随着企业向云端迁移以寻求灵活性与效率，网络边界逐渐模糊，传统的安全防护架构同样需要升级来应对新挑战。

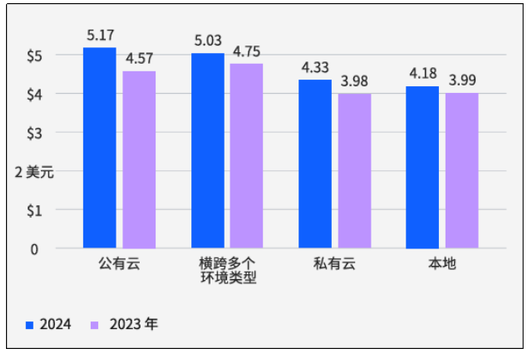


图 16 不同环境下的平均数据泄露成本（单位：百万美元）

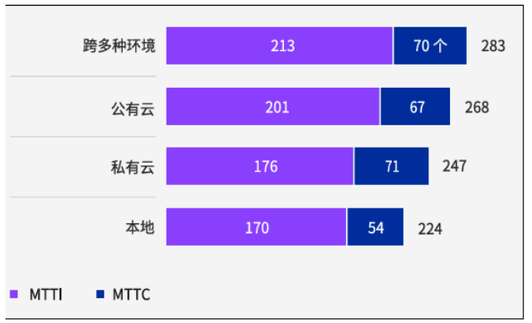


图 17 不同环境下的数据泄露事件响应时间（单位：天）

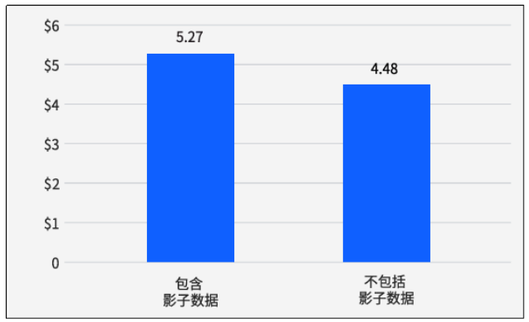


图 18 包含影子数据的泄露平均成本（单位：百万美元）

观点 7：影子数据的泄露将会带来更高的数据泄露成本，此类事件需要的响应时间也会更久。

尽管企业致力于扩展和优化数据管理策略，却往往忽视了一个重要方面——影子数据。当数据泄露事件涉及影子数据时，事件的持续时间会显著增加，造成的经济损失和后续处理成本也会更加高昂。

如图 18 所示，涉及影子数据的数据泄露事件的平均成本约为 527 万美元，比不涉及影子数据的数据泄露事件的平均成本高出 16.2%。影子数据虽然未直接暴露但可通过其他途径间接推断或重建的个人信息、公司机密等，其泄露可能带来更加广泛、深入的数据安全问题，因此造成的经济损失和声誉损害也更为严重。

如图 19 所示，较之不涉及影子数据的泄露事件，识别影子数据泄露平均需多用 26.2% 的时间，而遏制影子数据泄露平均需多用 20.2% 的时间。这些上涨的数据会造成平均生命周期达 291 天的数据泄露，而它比没有影子数据的数据泄露多出了 24.7%。

影子数据的复杂性和隐蔽性导致响应和遏制此类数据泄露事件需要更长的时间。数据泄露事件发生后，首先，识别和定位影子数据源是一个耗时且技术挑战高的过程。此外，影子数据可能涉及更广泛的个人信息和敏感数据，价值较大。对此类事件的影响评估、风险分析和应对措施的制定也需要更加谨慎和全面，也会增加响应时间和泄露成本。

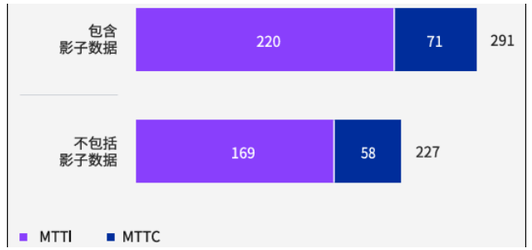


图 19 识别和遏制数据泄露（包括影子数据）所需的时间（单位：天）

云计算的快速发展也会加快影子数据的产生。首先，云服务的广泛使用和数据的频繁备份容易导致数据冗余和未授权数据的出现，形成影子数据。其次，随着远程办公的普及和企业快速向云端迁移，企业数据的管理变得更复杂，容易出现难以监管的影子数据。最后，云环境的动态性和多样性，使企业难以全面、实时地跟踪和管理所有数据，也为影子数据的形成提供了条件。报告指出，涉及影子数据的泄露事件中约有 75% 和云环境有关，另外的 25% 发生在本地环境，参考图 20。

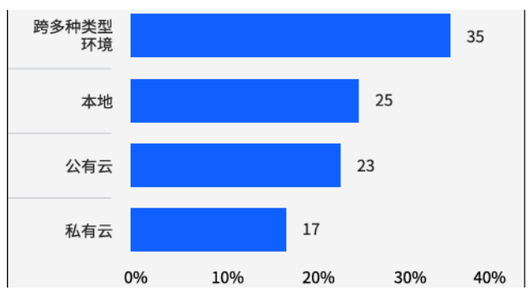


图 20 泄露事件中所含的影子数据的环境

3. 缓解措施

通过 2024 年 IBM 发布的数据泄露成本报告，我们总结了应对数据泄露事件的安全缓解措施，这些措施可以帮助以更低的成本、更短的时间识别并遏制泄露事件。

- 将安全融入软件开发和部署的每个阶段：企业应确保DevOps每个环节的安全性，严格遵循设计安全原则、开发安全原则和默认安全原则，而不是寄托于漏洞、安全事件发生后的修补和响应。
- 充分了解企业自身环境信息：大多数企业的资产和数据分布在多个环境中，可能涉及云环境和本地环境。因此，企业应尽可能地对自身的资产和数据进行全面监控和管理，让影子资产无处遁形。
- 应用AI和自动化安全技术：企业应在安全建设中应用AI和自动化安全技术，来提高对安全威胁事件的预警和响应速度以及准确性。另外，在选择AI相关的安全技术时，也需要关注AI技术本身的可靠性和成熟度，避免对企业安全系统的鲁棒性和稳定性造成影响。
- 进一步加强网络安全响应培训：企业应加强对企业员工网络安全应急响应能力的培训，提高企业检测、遏制和应对泄露事件的能力，降低数据泄露所造成的成本高昂且极具破坏性的影响，为运营连续性提供支持。

4. 总结

本文是对 IBM 发布的《2024 年数据泄露成本报告》进行的分析和解读，将报告全文的 14 个主题提炼为 7 个主要观点进行阐述、分析。通过报告可以看出全球数据泄露事件仍处于高发阶段，数据泄露成本仍处于上升趋势。虽然 AI 等自动化安全技术能够帮助企业降低数据泄露成本、缩短事件响应时间，但数据泄露问题依然十分严峻。因此，企业需要定期对自身的数据泄露风险进行评估，增加员工安全意识、技能培训，制定高效的事件应急响应与恢复措施，只有这样才能降低数据泄露事件发生的概率和危害。

绿盟科技创新研究院在云上风险发现和数据泄露领域已经开展了多年的研究。目前，我们已监测到数万个云端暴露资产存在未授权访问的情况。这些资产包括但不限于自建仓库（如源代码和镜像仓库）、公有云对象存储、云盘、OLAP/OLTP 数据库，以及各类存储中间件等。具体研究内容，可以参考《2023 公有云安全风险分析报告》[7]。

Fusion 数据泄露侦察平台是由绿盟科技创新研究院自主研发的一款创新产品，其能力如图 21 所示。Fusion 主要针对数据泄露测绘，集成探测、识别、泄露数据侦察于一体的平台，针对互联网中暴露的泛云组件进行测绘，识别组件关联的企业机构和组件风险的影响面，实现自动化的资产探测、风险发现、泄露数据分析、责任主体识别、数据泄露侦察全生命周期流程。

图 21 Fusion 核心能力

参考文献

[1]https://www.ibm.com/reports/data-breach.

[2]https://www.akamai.com/zh/glossary/what-is-attack-vector.

[3]https://www.jianshu.com/p/5eb78d6b1918.

[4]https://www.atlassian.com/zh/devops.

[5]https://www.threathunter.cn/blog/2024?categoryId=78595.

[6]https://baijiahao.baidu.com/s?id=1786005964624162935&wfr=spider&for=pc.

[7]https://www.nsfocus.com.cn/html/2024/92_0313/212.html.

源码&二进制组成成分分析现状研究

绿盟科技 创新研究院 王永吉

1. 软件组成成分分析主要背景

软件成分分析（SCA）旨在识别和管理软件项目中包含的开源组件，其中组件指的是重用的 TPL 及其对应的版本。基于 SCA 的结果，开发人员可以有效地跟踪软件项目的潜在威胁，如漏洞传播和许可证违规。

在检测目标中，目前主要包含两部分，一是源码，二是二进制文件。源码面对的主要场景为开发完成后对源码进行扫描，二进制文件为采购、引入的软件分析是否有安全危险。

2. 源码组成成分分析方法概述

基于源码的软件组成成分分析依赖两种分析方法：一是基于构建文件的软件组成成分分析方法，二是基于代码相似度的软件组成成分分析方法。

基于构建文件的软件组成成分分析方法，一般是检测安装过程的“安装痕迹”，即各个语言的构建文件（C 语言的 MakeFile 文件、Maven 项目的 pom.xml 文件等）分析得到被测项目使用了哪些第三方组件，因此该类方法主要是实现各类文件的分析算法提取部署组件库的名字以及版本，不需要大量的存储空间，耗费时间短，但是分析结果的准确性非常差，一是大部分安装痕迹非常模糊，对于很多正确的组件名以及版本无法正确地判断；二是很多组件，用户会选择性安装，而不是全部安装，比如仅在开发中使用的测试组件，requirement.txt 标记的 “extra == dev”，而工具无法判断用户是否真正安装这类组件，因此关联出大量漏洞、许可威胁，产生告警疲劳，使修复人员对于威胁疲于奔命。

基于代码相似度的组成成分分析方法需要两部分，一是相似度比对方法，二是各个组件版本代码特征库。对项目内代码与组建的特征库特征进行比较，若是相似度达到一定阈值，则认为使用了该组件。出于商业原因，笔者无法接触到各个 SCA 厂商的数据库，不能确定其所用的算法和特征抽取方法。但大部分组件的特征抽取采用基于目录以及摘要的默克尔树变体的方法 [1] 进行生成，在大规模组件库的背景下，能够进行快速计算、缩小存储空间。

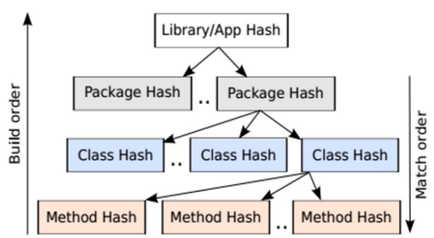


图 1 基于目录以及摘要的默克尔树变体的代码相似度识别方法

基于构建文件、代码相似度的软件组成成分分析方法适用领域各有不同，针对 Java、Python 等语言具有完备的构建过程以及开发管理仓库，大部分人的开发行为、流程应用到统一的构建过程，适用于基于构建文件的成分分析方法，对于组件分布广泛的 C/C++ 生态系统，开发人员往往手动导入组件，因此适用于基于代码相似度的组成成分分析方法。

基于构建文件、代码相似度的软件组成成分分析方法优缺点不同，基于构建文件的成分分析方法速度快、效率高、占用空间小、准性差、高度依赖于分析算法；基于代码相似度的软件组成成分分析算法即使只是抽取了代码的摘要，面对海量组件的情况下仍然占用空间大，因需要对海量组件进行匹配，分析速度慢，但是准确率高，高度依赖于摘要抽取算法，例如 oauthlib 组件为了修复 CVE-2022-36087 漏洞，仅仅修改了 IPv6addres 正则表达式中的 10 个字符升级成了新的版本，摘要抽取算法必须对目录、文件结构、变量等数据抽取得精准才能对组件的各个版本非常敏感。

3. 二进制组成成分分析方法背景

二进制软件在代码形成二进制文件的过程中会损失大量的特征，以 C++ 为例，在编译过程中 GCC 提供了 O_0~O_3 四种优化模式供开发者选择，随着优化级别的提升，代码损失的信息更多。

由一份相同的代码形成的二进制文件各不相同，编译器不同，不同的编译器可能会实现不同的优化策略，导致生成的二进制文件大小、执行速度或者其他性能指标有差异；编译时会为了安全、知识产权进行代码混淆，使用不同的代码混淆技术会导致生成的二进制文件结构和内容不同，如变量名混淆、控制流混淆、字符串加密等技术可以使得二进制完全不同；处理器架构也会影响最终的二进制文件，不同的处理器架构有不同的指令集和硬件特性，因此同一份源代码在不同架构的处理器上编译成的机器码会不同，例如，x86 架构和 ARM 架构的处理器使用不同的指令集，因此生成的二进制文件在这两种架构上可能是不兼容的，需要分别编译，生成的文件也会天差地别。编译器、代码混淆技术和处理器架构等因素会对生成的二进制文件产生影响，使得即使是相同的源代码，在不同条件下编译得到的结果也是不同的。

因此，由上所述，在编译过程中会损失大量信息，并且由于编译过程的区别，导致编译的二进制文件各不相同，使得基于源码的组成成分分析方法完全不适用于二进制文件。

在二进制软件成分分析中一般分析是使用反汇编工具将二进制代码反汇编成汇编代码进行代码相似性检测。有时候会进一步使用 IDA 把汇编代码转换为伪代码形式。但是一般还是使用汇编代码进行分析，对汇编代码进行相似性检测，由于上述提到二进制文件不同，因此反汇编的结果也不同，对于汇编代码相似性检测的难度类似于克隆代码第四种类型的相似性检测程度。

4. 二进制组成成分分析方法概述

二进制组成成分分析，是在二进制文件中识别可能存在的代码复制、代码克隆或者其他类型的代码重用情况，从而帮助发现潜在的安全风险或者违反知识产权的行为。二进制代码相似度（Binary Code Similarity Detection，BCSD），是二进制相似性分析的核心方法，用于检测两个或多个二进制代码文件之间的相似性，以此区分是否进行了组件复用。

目前针对软件供应链的二进制相似性检测方法尚不成熟，但是针对二进制相似性检测方法的历史由来已久，接下来分别介绍基于 SMT、基于 CFG 的同构程度、基于 I/O 相似性、基于深度学习的二进制相似性检测方法。

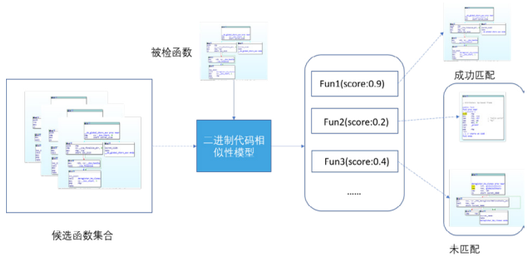


图 2 二进制文件相似性通用模型架构

基于 SMT 的二进制相似性检测方法，将代码语义转换为约束条件，通过约束求解器来判断代码的相似性。SMT 求解器是用于解决 SMT 问题的理论工具，它们能够自动分析给定的逻辑公式，并判断其是否可满足。在软件工程、硬件验证、形式化方法等领域，SMT 求解器被广泛应用于模型检测、程序分析、自动推理等任务中。David Y 等人 [2] 提出了一种基于 SMT 的统计方法，用于衡量二进制中两个过程之间的相似性。他们的相似性概念使得即使在使用不同编译器编译或被修改后，也能找到相似的代码。主要思想是利用组合相似性：将代码分解为较小的可比较片段，定义片段之间的语义相似性，并使用统计推理将片段相似性提升到过程之间的相似性。他们在一个名为 Esh 的工具中实现了这种方法，并应用于跨编译器和版本查找各种著名漏洞，包括 Heartbleed、Shellshock 和 Venom 等。他们展示了 Esh 产生了高准确性的结果，几乎没有误报的情况，足以说明该检测方法的准确度。

图 3 基于 SMT 的越是条件等价判断方法

基于 CFG 的二进制相似性检测方法，Karamitas C 等人实现了一种基于 CFG 的二进制相似性检测方法，用于研究两个程序之间的语法和语义差异。对于大型程序，二进制相似性检测可以通过函数匹配来实现，而函数匹配则被简化为比较程序的 CFGs（控制流图）和 / 或 CGs（调用图）之间的图同构问题。Karamitas C 等人 [3] 提出了一组从二进制的 CG 和 CFG 中精心选择的特征，这些特征可以被 BinDiff 算法变体使用，首先通过扫描唯一的完美匹配来构建一组初始的精确匹配，从而最小化误报，其次通过如最近邻策略来传播近似匹配信息。此外，Karamitas C 等人还研究了将马尔可夫聚类技术应用于函数 CFGs 的好处。在 Linux 内核(Intel64)、OpenSSH 服务器（Intel64）和 Firefox 的 xul.dll（IA-32）的各个版本上进行了一系列实验来评估提出的函数特征，并将其原型系统与当前最先进的二进制相似性检测软件 Diaphora（基于上文描述的 SMT）进行了比较。

图 4 基于控制流图的二进制相似度比对方法

基于 I/O 行为的二进制相似性检测方法，Jannik Pewny 等人 [4] 提出了一种基于 I/O 行为的二进制漏洞搜索方法，旨在识别已知漏洞的漏洞签名，并利用这些签名部署在不同 CPU 架构（如 x86 与 MIPS）上的二进制文件中查找漏洞。针对不同 CPU 架构的多样性，该方法面临着诸多挑战，如 CPU 模型之间指令集体系结构的不可比性。为了解决这一问题，该方法首先将二进制代码转换为

中间表示形式，生成带有输入和输出变量的赋值公式。其次，对基本块进行具体输入的采样，观察其 I/O 行为以把握其语义。最后，利用 I/O 行为找到行为与漏洞签名相似的代码部分，有效地揭示包含漏洞的代码部分。最终设计并实现了一个用于在可执行文件中跨架构查找漏洞的工具原型。该原型目前支持三种指令集架构（x86、ARM 和 MIPS），并可以在任何这些架构的漏洞二进制代码中发现漏洞。他们展示了无论底层软件指令集如何，都能够发现 Heartbleed 漏洞。

基于人工智能的二进制相似性检测方法。基于人工智能的检测方法主要包含两种解决方式；代码嵌入和图嵌入。

代码嵌入（Code embeddings），许多研究人员尝试采用现有的自然语言处理（NLP）技术，通过将汇编代码视为文本来解决二进制函数相似性问题。这些解决方案处理标记流（如指令、助记符、操作数、标准化指令）并输出每个代码块一个嵌入、每个指令一个嵌入或两者都输出，OrderMatters[5]、Trex[6] 等模型均用的此类方法。

图嵌入（Graph embeddings），其他研究建立在计算图嵌入的机器学习方法的基础上，这些非常适合捕获基于函数控制流图的特征，这些特征本质上是跨架构的。这些嵌入可以通过自定义算法或更复杂的机器学习技术生成，Xiaojun Xu、Zeping Yu、Rui Cao 等人 [7][8][9] 的研究均是基于图嵌入方法。

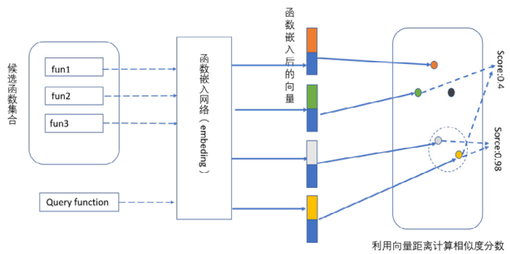


图 5 基于人工智能的二进制相似性识别框架

5. 总结

目前的二进制检测方法中，基于 SMT 的检测方法在实际操作中，约束求解中的理论证明仍然存在很多问题和挑战，同时该方法计算复杂度很高，代价很大。I/O 操作和全局 / 堆内存访问中遇到的值在编译器很难优化这些行为，可以利用这个特征进行相似性比对任务，效果较好。基于人工智能的二进制相似性比对算法近几年突飞猛进，虽然分为代码嵌入和图嵌入两个流派，但是在检测粒度（基本块、指令）、跨架构、跨优化等级、识别函数语义都有了非常大的进步。

传统二进制相似性方法与软件供应链二进制组成成分中的相似性检测方法仍有区别。软件供应链中二进制软件成分分析的两大特点，一是组件库量级大，组件种类多、组件内版本数量大、差异小；二是组件在二进制相互包含、相互依赖，检测结果多样化。针对 SMT 的二进制相似性方法完全不适用于软件供应链领域，复杂度、代价很大，无法面对检测目标众多的软件供应链领域；基于 I/O 的行为的二进制相似性检测方法，无法区分一个组件中的多个版本，例如 oauthlib 组件修改了 IPv6addres 正则表达式中的 10 个字符升级成了新的版本；基于人工智能的检测方法，数据集过大，无法区分二进制内部组件之间的依赖关系。

目前，软件供应链中对软件组成成分的分析方法仍然局限于分析和识别二进制文件安装后释放的动态链接库等文件的名称和哈希值等简单特征。因此，在面对软件供应链领域时，我们仍然需要寻找一种新的二进制相似性分析方法。或许，大型语言模型将成为最终的解决方案。

参考文献

[1]Libscout:Reliable Third-Party Library Detection in Android and its Security Applications CCS ’ 16, October 24–28, 2016, Vienna, Austria

[2]David Y, Partush N, Yahav E. Statistical similarity of binaries[J]. Acm sigplan notices, 2016, 51(6): 266-280.

[3]Karamitas C, Kehagias A. Efficient features for function matching between binary executables[C]//2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER). IEEE, 2018: 335-345.

[4]Pewny J, Garmany B, Gawlik R, et al. Cross-architecture bug search in binary executables[C]//2015 IEEE Symposium on Security and Privacy. IEEE, 2015: 709-724.

[5]Yu Z, Cao R, Tang Q, et al. Order matters: Semantic-aware neural networks for binary code similarity detection[C]//Proceedings of the AAAI conference on artificial intelligence. 2020, 34(01): 1145-1152.

[6]Pei K, Xuan Z, Yang J, et al. Trex: Learning execution semantics from micro-traces for binary similarity[J]. arXiv preprint arXiv:2012.08680, 2020.

[7]Xu X, Liu C, Feng Q, et al. Neural network-based graph embedding for cross-platform binary code similarity detection[C]//Proceedings of the 2017 ACM SIGSAC conference on computer and communications security. 2017: 363-376.

[8]Yu Z, Zheng W, Wang J, et al. Codecmr: Cross-modal retrieval for function-level binary source code matching[J]. Advances in Neural Information Processing Systems, 2020, 33: 3872-3883.

[9]Yu Z, Cao R, Tang Q, et al. Order matters: Semantic-aware neural networks for binary code similarity detection[C]//Proceedings of the AAAI conference on artificial intelligence. 2020, 34(01): 1145-1152.1111

网络安全政策导读（2024年9—12月）

绿盟科技 总体技术部 林涛

栏目说明：

本专栏基于绿盟科技团队在网络安全政策法规方面的日常跟踪，筛选国内外当期热点政策法规文件，并重点结合网络安全产业发展，对其内容和影响等进行简要分析。本期研究的国内外政策法规的发布时间范围为 2024 年 9—12 月。

更多内容敬请关注“网络安全罗盘”和“绿盟科技”微信公众号。

1. 国内篇

1.1《关于促进数据产业高质量发展的指导意见（征求意见稿）》发布

【内容概述】9 月 27 日国家数据局发布。《意见》在数据安全方面提出要提高数据领域动态安全保障能力。一要创新数据安全产品服务，推动基础设施安全、数据安全、应用安全协同发展，加强身份认证、数据加密、安全传输、合规检测等技术创新，培育壮大适应数据流通特征和人工智能应用的安全服务业态。支持企业创新数据分类分级、隐私保护、安全监测、应急处置等数据安全产品和服务。二要加强动态数据安全保障，扩大数据空间、区块链、隐私计算等可信流通技术及模式应用范围，增强数据可信、可控、可计量开发利用能力。建立健全数据安全风险识别、监测预警、应急处置等相关规范，落实数据流通利用全过程相关主体的安全责任。健全数据分类分级标准，加强对涉及国家安全、商业秘密、个人隐私等数据的保护。

【专家解读】

《意见》是国家数据局成立以来，提出的首份关于数据产业发展的专项政策文件。

2023 年初，工业和信息化部等十六部门联合印发了《关于促进数据安全产业发展的指导意见》（以下简称《数安产业意见》），对于数据安全产业发展也提出了指导。两个文件可结合来看。

从产业体系角度来看，《意见》明确了三个重要问题。一是产业发展目标规模，给出了 15% 的年复合增长率目标，该目标明显低于《数安产业意见》所提出的 30% 年复合增长率目标（2025 年实现 1500 亿元产业规模），这可能与数据产业基数更大有关；而至于具体产业规模量化目标，《意见》并未提及。二是产业门类划分，《意见》将数据产业具体划分为数据资源、数据技术、数据服务、数据应用、数据安全、数据基础设施 6 个细分门类，并对其中的数据安全产业细分了产品应用、技术创新、服务 3 个子门类；这与《数安产业意见》所提出的“创新、服务、标准、应用”基本一致。三是对于产业的定位，《意见》将数据安全产业发展定位为“提高数据领域动态安全保障能力”，并明确了“动态安全”的主要范畴，包括可信流通技术和模式，风险识别、监测和应急等，可见在《数安产业意见》基础上有创新发展。

《意见》目前征求意见已结束，诸如对数安产业的监管分工、监管模式、推进举措等具体落地，值得安全行业重点关注。

1.2《网络数据安全管理条例》正式发布

【内容概述】9 月 30 日国务院第 40 次常务会议通过。《条例》共 9 章 64 条，主要规定了以下内容。一是提出网络数据安全管理总体要求和一般规定；二是细化个人信息保护规定，明确处理个人信息的规则和应当遵守的具体规定；三是完善重要数据安全制度；四是优化网络数据跨境安全管理规定；五是明确网络平台服务提供者义务。

【专家解读】

《条例》先前于 2021 年公开征求社会意见，在等待三年漫长修改完善后终于获得审议通过。

与征求意见稿相比，一个最大的特点就是对原有的机制建设部分做了大量删减和调整，主要包括数据安全应急机制、数据安全审查制度、数据出口管制和反制机制、数据交易管理制度、数据安全审计制度 5 项制度和机制。究其原因，大概存在三方面考量，一是制度的吸收，如数据应急机制吸收到新的数据协调机制中作为一项职责；二是制度的取代，如数据安全审计制度，被后来建立的个人信息保护合规审计、重要数据风险评估、重要数据出境安全评估等取代替换；三是职责的迁移，如对于数据交易管理制度，相关内容属新成立的国家数据局统一管理。

当然，《条例》也突出强调了一些新规定和机制，如国家数据安全工作协调机制、出境安全专项工作机制、网络数据损害赔偿责任保险机制、个人信息保护社会责任报告机制 4 种机制和制度。

就条例后续的落实而言，或有四个方面需重点关注。一是关于重要数据的界定，条例只保留定性的规定，具体数据范围将由各级各行业的“重要数据目录”具体划定。二是主管部门在检查发现较大风险时的处理权限和程序规定如何细化完善。三是新增的管理机制，如数据出境安全专项工作机制等的范围和职责有哪些。四是新增重要应用机制，如网络数据损害赔偿责任险具体内容和规定有哪些等。

1.3《关于加快公共数据资源开发利用的意见》发布

【内容概述】10 月 9 日中共中央办公厅、国务院办公厅发布。《意见》旨在加快公共数据资源开发利用，充分释放公共数据要素潜能，推动高质量发展，从扩大资源供给、规范授权运营、鼓励应用创新、营造良好环境、强化组织保障等方面提出了 17 项具体措施。《意见》指出应加强安全管理，强化数据安全和个人信息保护，建立健全分类分级、风险评估、监测预警、应急处置等工作体系，开展公共数据利用的安全风险评估和应用业务规范性审查。

【专家解读】

数据资源开发利用是《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》（“数据二十条”）所明确的国家数据基础制度建设的重要基础。《意见》立足公共数据资源的基础性、战略性属性，为促进公共数据合规高效流通使用提出明确指引。

近期国家数据管理部门密集发布系列政策规范，如《关于促进企业数据资源开发利用的意见》《国家数据标准体系建设指南》等。与这些政策规范相比，《意见》属于中办发文，发文层级更高，更凸显中央对公共数据开发利用工作的关注。

从安全角度看，《意见》对公共数据资源开发利用工作中的安全也提出了关注重点，包括四个方面：一是强化全程监管，将数据资源的生产、加工使用、产品经营等关键环节纳入监管范畴；二是加强审查和评估，持续推进分类分级、风险评估、监测预警、应急处置等工作体系建设；三是强化运营机构的主体责任，加强风险识别和管控；四是严格保密管理，坚守保密红线。

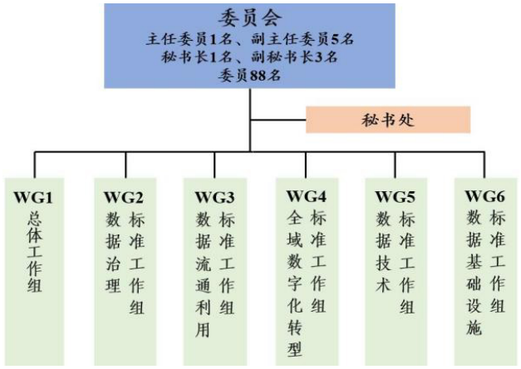
1.4 全国数据标准化技术委员会成立大会暨第一次全体委员会议在京召开

【内容概述】10 月 28 日全国数据标准化技术委员会成立大会暨第一次全体委员会议在京召开。会议审议通过了全国数据标准化技术委员会章程、秘书处工作细则、标准制修订工作程序等制度文件，以及全国数标委 2024—2025 年工作要点、下设工作组组成方案。

【专家解读】

继《国家数据标准体系建设指南》发布不久，此次全国数据标准化技术委员会（以下简称数标委）的成立，是国家数据局在数据标准化建设领域的又一重要动作。从国家数据局此前发布的制度体系来看，数据标准是数据基础设施的重要组成部分。而数标委的成立，则是系统推进数据标准化工作的主要平台和力量。

从推进机制来看，数标委由 98 名委员，6 个工作组（WG）构成（如下图所示）。通过设立专门的标准化技术委员会（TC）来推进领域和行业的标准化工作，是目前较为通行的做法。我国在国家标准委（SAC）下，设有各领域的 TC 有 609 个，具有数据相关标准职能的 TC 约有 37 个，而本次成立的数据标准化委员会(SAC/TC609）其体系性更强。



从职责内容来看，数安标委“主要负责数据资源、数据技术、数据流通、智慧城市、数字化转型等基础通用标准，支撑数据流通利用的数据基础设施标准，以及保障数据流通利用的安全标准等”，基本与《国家数据标准体系建设指南》保持一致。且在数据安全领域，其侧重点在于数据流通利用相关的安全保障标准。日前，数标委已公布了由 37 项标准构成的《2024—2025 年拟制修订的重点标准项目清单》。

从行业影响来看，数标委作为数据领域标准化工作的统一平台，其成员单位无疑在标准信息动态、规划布局、技术资源等方面具有一定发展先机。目前从 98 家委员单位人员构成上看，网络和数据安全厂商数量不多，且数标委下设的 6 个工作组成员尚未完全明确，这或为数据安全行业预留了一定竞争空间。

1.5 商用密码检测机构（商用密码应用安全性评估业务）资质申请通过技术评审的机构名单公示

【内容概述】10 月 28 日国家密码管理局官网公布商用密码检测机构（商用密码应用安全性评估业务）资质申请通过技术评审的机构名单。名单共 112 家机构，其中北京最多 24 家，其次是广东 8 家，重庆 7 家，天津、浙江、四川分别为 6 家，吉林、甘肃、宁夏、

青海、贵州各 1 家，新疆、西藏、黑龙江暂无。

【专家解读】

这是 2023 年《商用密码检测机构管理办法》发布实施以来，国家密码局公布的首批通过评审的密评机构公示名单，也是国家将密码产品检测机构和密评机构纳入统一管理之后的首次密评机构公示名单。

根据《商用密码检测机构管理办法》，机构申请成为密评检测机构，需要经过形式审查、实质审查等主要流程，国家密码局认为需要时还要通过专业机构开展的技术评审。机构获得资质证书后，检测机构还需履行数据留存、年度报告等主要义务；且在日常开展业务过程中，机构需接受各级密码管理部门的监督检查。目前，认定检测机构的主要操作依据是《商用密码检测机构资质认定基本规范》。

作为密码应用领域的重要监管抓手之一，国家密评工作的推进进展情况值得网络安全行业重点关注。一是关注密评适用范围的扩大，根据《密码管理条例》规定，密评的范围不再限于《密码法》规定的“关基”，而是扩大为“法律、行政法规和国家有关规定要求使用商用密码进行保护的网络与信息系统”，这进一步扩展了密评的业务范围。二是密切推动密评相关业务拓展，应立足产品技术服务创新，加强对监管侧业务的跟进，除了监管部门之外，检测机构也是重要的服务对象之一，重点在于相关技术产品的持续性发展、供应链保障等方面。三是密评和等保测评的关联，从名单构成来看，目前有接近 70% 的密评检测机构同时还是等保测评机构，因此相关业务线条之间的统筹协调或可成为企业、行业业务布局时的一个重要考量因素。

1.6《可信数据空间发展行动计划（2024—2028 年）》发布

【内容概述】11 月 23 日国家数据局发布。《行动计划》主要包括三大行动：一是实施可信数据空间能力建设行动，通过构建可信管控能力，提高资源交互能力，强化价值共创能力，打造可信数据空间的核心能力体系。二是开展可信数据空间培育推广行动，主要是布局企业、行业、城市、个人、跨境五类可信数据空间建设和应用推广，探索各类数据空间的场景创新、模式创新、机制创新。三是推进可信数据空间筑基行动，围绕制定关键标准、攻关核心技术、完善基础服务、强化规范管理、拓展国际合作五个方面，全面夯实可信数据空间发展基础。

【专家解读】

可信数据空间相关制度是国家数据局提出的数据要素系列基础制度之一。

从定位看，可信数据空间归属于数据基础设施中的数据流通基础设施细分门类，而在该细分门类中，除了数据空间，还包括区块链、高速数据网等为代表的数据流通基础设施。（根据国家数据局相关制度设计，数据基础设施涵盖网络基础设施、算力基础设施、流通基础设施、安全基础设施四个细分门类）

从载体形式看，《行动计划》明确了可信数据空间所需具备的三类核心能力：可信管控、资源交互、价值创造。从中不难发现，可信数据空间在具体载体形式上，涵盖而又高于传统的数据中心，其更兼具了数据交易、数据安全的某些支撑保障功能。进一步看，后续可信数据空间建设或将存在新建以及对存量数据中心升级改造两种可能的推进模式。

从类型来看，《行动计划》以“培育推广行动”专章指明了可信数据空间的主要类型，即企业可信数据空间、行业可信数据空间、城市可信数据空间、跨境可信数据空间等。这些不同类型的数据空间，在建设定位、辐射范围、核心价值等方面都有较明显区别。

1.7《电力监控系统安全防护规定》

【内容概述】12 月 11 日国家能源局发布。《规定》包含七个主要章节：一是明确电力监控系统范围，将罗列典型系统名称改为列举功能。二是优化安全分区要求，在坚持原分区策略的基础上，调整原《规定》阐述逻辑及表述。三是强化安全接入区防护要求，明确安全接入区加密认证、安全监测等技术要求。四是强化技术防护措施，在坚持十六字原则的基础上，补充安全免疫、态势感知、动态评估和备用应急措施。五是定义电力监控专用网络，明确承载电力监视和控制业务的专用广域数据网络、专用局域网络以及专用通信线路属于电力监控专用网络范畴。六是强化供应链及电力监控系统专用安全产品管理，明确运营者应当以合同条款的方式对电力监控系统供应商提出安全要求，明确由国家电力调度控制中心牵头组建电力监控系统专用安全产品管理委员会。七是优化技术监督管理，明确不同主体技术监督的工作要求，增加技术监督过程中风险管控措施。八是细化罚则。

【专家解读】

国家能源局在 2024 年 7 月发布了《征求意见稿》，从最终稿对比来看，有以下两方面内容可着重学习理解。

首先是内容的三大调整变化。一是技术要求的补充完善。大幅加强了“安全技术”章节内容，对于安全分区、边界要求、设备产品要求等做出较多、较细化的补充和调整。二是管理机制的优化完善，尤其是将“电力监控系统专用安全产品管理工作机制”，改为“电力监控系统专用安全产品管理委员会”，并对委员会构成、工作

机制等做出了对应修改，反映了务实和开放的管理思路变化。三是对关键名词解释进行界定，尤其是对事关电力系统安全的“电力监控系统专用安全产品”“电力监控系统”进行了较多细化和明确。

其次是行业两个重要关注点。《规定》是电力行业网络安全的重要管理依据和实践规则，其内容要求不仅明确了监管方向，也为网络安全供给指明了市场。一是从监管抓手来看，《电力监控系统专用安全产品目录》仍然值得重点关注，尤其相关配套规范标准的修订等工作。二是从市场机会来看，《规定》所明确提出的分区管理、边界隔离、监测预警、加密认证，以及专用安全产品等，或均会成为触发未来行情的新增长点。

2. 国外篇

2.1 美国白宫召开人工智能基础设施领导力圆桌会议（White House Roundtable on U.S. Leadership in AI Infrastructure）

【内容概述】

9 月 12 日美国白宫发布。美国白宫科技政策办公室（OSTP）召集人工智能企业、研究机构、政府机构、数据中心等相关实体，召开圆桌会议，并宣布一系列举措，加强美国先进人工智能运营所需的基础设施。举措包括：白宫成立人工智能数据中心基础设施工作组；扩大对联邦、州和地方当局数据处理中心的许可；能源部牵头组建人工智能数据中心团队，支持数据中心建设；能源部将继续开发和共享资源，将重新利用已关闭的煤电厂；美国陆军工程兵团将加快关键领域合格 AI 数据中心建设；政府加强与行业领军企业合作；行业企业同意努力实现净零排放目标。

【专家解读】

本次会议隶属于拜登政府“负责任创新综合战略”（comprehensive strategy for responsible innovation）的范畴，旨在保持美国在人工智能领域的领导地位。

会议的核心议题是采取措施，建立健全高级 AI 运营所需的大规模 AI 数据中心和电力基础设施。会议公报所提出的举措涉及组织机构(成立 AI 数据中心基础设施工作组)、管理机制（加快数据中心许可效率和资金保障)、电力保障（能源部数据中心清洁能源计划）等多个方面。

本次会议的召开，进一步拓展了美国发展人工智能举措的广度和深度。从广度上看，不仅注重安全，更加强调发展，与此前美国政府发布的《关于安全、可靠和可信任地开发和使用人工智能的行政令》等安全类政策法规形成互补。从深度上看，将人工智能的发展目标细化到人工智能基础设施领域，并分解为机构、机制、资源等具体举措，强调人工智能发展的关键基础环节。

当前，我国也在持续推进人工智能发展及其监管，美国所采取的综合式举措或将对完善我们相关立法和政策体系具有一定借鉴意义。另外，我国正在大力开展的新基建、数据要素建设等工作，尤其是包含算力中心等在内的信息基础设施建设工作，对人工智能发展无疑具有重要战略意义；而如何加强此类基础设施的网络安全和数据安全保障能力，则是值得网络安全行业关注的重要市场机会。

2.2 美国国会表决通过 2025 年国防授权法案，网络安全预算略减

【内容概览】

2024 年 12 月 18 日美国参议院通过。旨在强化美军的全球竞争力。法案重点投资于先进技术和研发，包括量子信息、人工智能和网络空间安全，以应对新兴威胁并推动技术创新。此外，法案支持提升国防工业基础，包括建立先进制造工厂和增强潜艇及造船能力。法案还强调了对军人福利的支持，以及加强与盟友和伙伴国的安全合作，确保国际和平与稳定。

【专家解读】

法案最初于 2024 年 4 月 18 日在众议院提出。后历经讨论修改，并于上周三（12 月 11 日）在众议院表决中以 281:140 通过，本次在参议院表决中则以 85:41 通过。后续将提交总统签署后正式生效。按照法案文本，2025 年度国防授权总预算为 8950 亿美元，较 2024 年增加 90 亿美元，增幅约 1.01%。据初步统计，授权法案中网络安全的预算约为 14.4 亿美元，与 2024 年的 14.5 亿美元相比略减 0.1 亿美元，降幅为 0.69%。从网络安全预算的细分领域来看，2025 年美国网络安全国防预算支出靠前的领域主要有：能源网络安全和技术、军队作战和预备役网络安全、网络弹性和网络安全政策支持、网络安全教育等。

聚焦法律法规，洞察数据安全内涵

绿盟科技 创新研究院 高翔

摘要：数据安全的目的之一是保护敏感信息不受黑客攻击后遭泄露、篡改或破坏，数据安全相关法规和标准共同帮助实现这一目标。本文分析了国内外数据安全法律法规，从合规视角出发，介绍了涵盖数据要素全生命周期的数据安全新内涵。

关键词：数据安全 法律法规解读 GDPR 网络安全法

1. 数据安全法律法规与标准概述

1.1 全球数据安全立法现状

根据联合国贸易和发展组织（UNCTAD）统计，全球 77% 的国家（共 194 个国家）完成了数据安全和隐私立法或者已经提出法律草案，其中包括欧盟、美国、中国、俄罗斯、印度、澳大利亚、加拿大和日本等绝大多数国家。随着数字化转型的不断推进与深入，数据安全与隐私问题越来越严峻，现代化的数据安全与隐私保护立法已成为全球趋势。

2018 年 5 月 25 日，欧盟正式实施《通用数据保护条例》（GDPR）用以保护欧盟成员国境内企业的个人数据，也包括欧盟境外企业处理欧盟公民的个人数据以及公民享有的各项数据权利。

受 GDPR 的广泛影响，全球其他国家也陆续推出了各自相关的法规：巴西于 2019 年 7 月通过《通用数据保护法》（葡萄牙语简称 LGPD）的最终版本，于 2021 年 5 月生效；印度于 2018 年 12 月公布修改后的《2019 年个人数据保护法（草案）》（Personal Data Protection Bill, 2019）；泰国于 2020 年 5 月正式实施《个人数据保护法》（Personal Data Protection Act，PDPA）等。同样深受 GDPR 影响，美国各个州在数据隐私领域上纷纷重新立法，包括加利福尼亚州（加州）、蒙佛特州、夏威夷、马里兰、马萨诸塞、密西西比和华盛顿等。其中，最具代表性的是加州于 2018 年 6 月通过的《加州消费者隐私保护法》（California Consumer Privacy Act，CCPA）由于影响涉及大部分知名 IT 科技公司，如惠普、Oracle、Apple、Google 和 Facebook 等，该方案从立法到颁布备受各界人士的关注。2019 年 10 月，美国加州州长正式签署 CCPA 的最终法案，已于 2020 年 1 月 1 日正式生效。CCPA 与 GDPR 类似，同样对企业提出更高的数据合规性要求，据 IAPP 和 OneTrust 调查结果显示，大约仅有 2% 的受访者认为他们的企业已经完全做好了应对 CCPA 的准备。

在执法方面，欧盟 GDPR 相较其他国家的法规，已经进入全面执法阶段，其典型的趋势是多数成员国已经陆续开出多张违反 GDPR 的罚单。其中，英国执法力度最大，英国 ICO（Information Commissioner's Office）于 2019 年 7 月分别对英国航空公司和万豪国际集团数据泄露事件分别开出 1.83 亿英镑和 9900 万英镑的巨额罚单。此外，Google 罚款事件非常具有代表性，备受关注——作为一家大型国际互联网公司，Google 却已被欧盟的两个国家罚款：2019 年 1 月被法国处罚 5000 万欧元，原因是执法方认为 Google 的隐私条款未充分体现 GDPR 公开透明和清晰原则；2020 年 3 月被瑞典处罚 700 万欧元，原因是 Google 未充分履行 GDPR 赋予用户的数据“遗忘权”。GDPR 立法与执法的严苛程度，从以上事件可见一斑。

1.2 国外数据安全法律法规

（1）欧盟 GDPR

欧盟于 2018 年 5 月 25 日正式实施了《通用数据保护条例》（General Data Protection Regulation,GDPR），是一项保护欧盟公民个人隐私和数据的法律，其适用范围包括欧盟成员国境内企业的个人数据，也包括欧盟境外企业处理欧盟公民的个人数据。

GDPR 由 11 章 99 个条款组成，是一项“大而全”的个人数据保护框架。

第 83 条给出违反的罚款额度，即最高可处罚 2000 万欧元的行政罚款，或对企业以最高占上一财政年度全球总营业额 4% 的行政罚款，取两者最高值。

GDPR 赋予了用户（数据主体）知情权、访问权、修正权、删除权（被遗忘权）、限制处理权（反对权）、可携带权、拒绝权 7 项基本权利，其中删除权是一项引人注目的用户“特权”之一。即在第 17 条中，在个人数据已不再是数据控制者和处理者的收集和处理目的等 6 种情况下，赋予了用户删除权。

官网提供了两个有趣的例子 [1]。

Example 1：假如你加入了一个社交网站，但过了一段时间，你决定离开这个社交网站。那么你可以行使“被遗忘权”，即要求公司删除属于你的个人数据。

Example 2：当你用你的名字使用搜索引擎搜索时，出现一个很久以前，与你有关的债务偿还协议，但现在已无关紧要。对于一

个普通人，他有权利要求删除这些网络信息。用户除了以上的权利，数据主体还拥有知情权、访问权、修正权、限制处理权（反对权）、可携带权、拒绝权等基本权利。但笔者认为，“被遗忘权”是 GDPR 赋予用户非常大的一项权利。我们每个人平均一年要注册十多个网站 /APP，但很多网站 /APP 是低频访问或者后续一直不用的状态。用户的个人数据却被互联网公司收集存储和处理，用户感觉到“个人数据扩散不可控”。如实说，笔者对一些公司网站的“注册”有种恐惧感，例如个人数据被贩卖给第三方从而收到骚扰电话或广告邮件的轰炸、另外一些“小网站”被黑客攻击造成数据泄露的概率也更高。若这些网站 /APP 提供一键删除注册信息的功能，作为用户，肯定乐于行使该项权利。

（2）美国 CCPA

美国已有多个州先在数据安全与隐私保护进行了立法，其中最著名的要数 2018 年 6 月加州通过《加州消费者隐私法案》（California Consumer Privacy Act, 简称 CCPA）[2]。该法案被称为美国“最严厉和最全面的个人隐私保护法案”，于 2020 年 1 月 1 日生效。

① “个人信息”的定义

CCPA：“个人信息”系指直接或间接地识别、关系到、描述、能够相关联或可合理地联结到特定消费者或家庭的信息。

解读：CCPA 称为“个人信息”，其实和“个人数据”是同一个概念。受 GDPR 的影响，CCPA 同样采用了类似宽泛的定义。根据定义，罗列出了 11 种个人信息类别，包括姓名、驾照等信息，也有互联网 IP 标识符、标识符。有特点的是，把反映消费者偏好、特征、心理倾向、偏好、倾向、行为、态度、智力、能力和资质的画像也列入了个人信息范畴。比 GDPR 更加广泛的“个人信息”范畴给企业个人敏感数据的梳理、治理带来了巨大的工作量和挑战。

② 用户如何行使 CCPA 赋予的“访问权”

CCPA：在 CCPA 的 1798.100 条款中，企业从可验证消费者处收到要求访问个人信息的请求后，应立即采取措施向消费者免费披露和提供本节所要求的个人信息。个人信息的提供可通过信件或电子方式，如果以电子方式提供，信息应以便携方式提供并且在技术可行限度内采用易于使用的形式。

解读：CCPA 也赋予了消费者知情权、访问权、删除权、限制处理权和拒绝权等权利。CCPA 访问权的特色在于回复的“及时性”、反馈的“便携式”——邮件发送等形式，这比 GDPR 赋予的“访问权”更加具体。当用户需要查看或确认采集信息，无疑这条法规提供了极大的便利。但反过来说，给企业造成了一定负担。

③ CCPA 如何惩罚违规的企业

CCPA: 在 1798.155 条款中，对于法规企业，每次违规行为可能要承担高达 7500 美元的民事罚款。另外，在 1798.145 中，对于违规企业，为每个消费者每次事件赔偿不少于 100 美元且不超过 750 美元的赔偿金，以数额较大者为准。

解读：罚款最有特色的地方，考虑到消费者的损失且量化为影响的消费者数量，将赔偿每一个消费者 100 ～ 750 美元的赔偿金，这与 GDPR 的罚款机制不同。若一个企业违规涉及的消费者信息有 100 万条，那么它至少要承担 1 亿美元的罚款。

1.3 中国数据安全法律法规

我国于 2017 年 6 月 1 日正式实施《中华人民共和国网络安全法》（以下简称《网络安全法》）。它是我国首部较为全面规范网络空间安全管理方面问题的基础性法律，不仅包括网络运行安全、关键信息基础设施的运行安全，同时给出数据安全与个人信息保护的基本规定。

自 2019 以来，我国数据安全相关立法进程明显加快：根据《网络安全法》，国家互联网信息办公室（以下简称“网信办”）分别于 2019 年 5 月和 6 月发布了《数据安全管理办法（征求意见稿）》和《个人信息出境安全评估办法（征求意见稿）》等法规；同年 10 月 1 日网信办正式实施《儿童个人信息网络保护规定》，对儿童个人信息安全进行特殊和更加严格的保护。2020 年 5 月，我国发布《中华人民共和国民法典》，于 2021 年 1 月 1 日起实施，其首次在我国法律中明确且具体提出“隐私权”的概念，并确立隐私权范围和个人信息保护的一些基本规范。2020 年 7 月，我国对外发布《中华人民共和国数据安全法（草案）》[以下简称《数据安全法（草案）》]，确立了数据分级分类保护、数据安全风险评估、应急处置机制和安全审查的重要制度，明确了开展数据活动必须履行数据安全保护义务等内容。2020 年 10 月，《中华人民共和国个人信息保护法

（草案）》[以下简称《个人信息保护法（草案）》] 在人大网公开，该法律赋予必要的域外适用效力，以充分保护我国境内公民的权益，同时回应 GDPR 等国外法规的同等效力；完善和丰富了个人各项数据权利，赋予个人包括知情权、决定权、查询权、更正权、删除权等；相比《网络安全法》，对违法行为加大了惩处力度。《数据安全法》和《个人信息保护法》作为两部较为综合性的法律，前者更加强调总体国家安全观，对国家利益、公共利益和个人、组织合法权益方面给予全面保护；后者则更加侧重于对个人信息、隐私等涉及公民自身安全的个人信息与权益的保护。

在跨境数据传输上，网信办于 2022 年先后公布了《数据出境安全评估办法》以下简称评估办法和《个人信息出境标准合同办法》，至 2023 年 3 月 1 日《评估办法》规定的整改期届满。据公开资料的不完全统计，截至 2023 年底已公布通过评估的企业有 29 家，与国家网信办和各地省级网信办受理的千余件申报相比，数据出境申报通过率或仅为百分之一。多数存在数据出境业务、按规定需进行评估的企业仍然在评估流程中，而《评估办法》规定评估结果有效期仅为两年，企业面临反复而冗长的评估，可能对跨境业务造成不利影响。而 2023 年 9 月 28 日网信办公布的《规范和促进数据跨境流动的规定（征求意见稿）》，在上述规定的基础上进一步细化规范，适当优化了审批程序和适用范围，一定程度地明确和放松了监管要求。此外，在 2023 年下半年，粤港澳大湾区范围内出台了一系列数据跨境相关的地方法规和标准指南，提供了规范数据跨境活动的多种机制，走在了数据跨境规范管理落地实践的前列。

在数据互联互通与数据交易上，中共中央国务院于 2022 年底发布《关于构建数据基础制度更好发挥数据要素作用的意见》、2023 年 2 月发布《数字中国建设整体布局规划》，国家数据局于 2023 年 10 月 25 日的正式成立标志着国家数字化发展迈出关键一步。该机构的宗旨是整合数据要素资源、在各个领域驱动数字创新引领的技术进步与市场发展。而在 2023 年的最后一天，国家数据局会同网信办、工信部、科技部等部门联合印发《“数据要素 ×”三年行动计划（2024—2026 年）》，是其自成立以来的首次重磅发声。

2. 合规视角下的数据安全新内涵

随着相关法规的逐步实施与完善，数据安全问题已经不再仅仅是企业“关起门”来处理的内部问题，同时也是国家监管部门参与进来的监管问题。企业的数据安全建设在满足内部数据安全需求的同时，也必须遵循法律法规的合规性条款。现阶段，对企业数据安全提出了哪些新的要求？我们从如下几个方面进行分析：

（1）需要保护的数据范围更大：原先企业以重要资产的数据安全防护为视角，重点保护的数据对象是企业敏感数据，同时也包括一小部分个人隐私数据，比如用户的登录密码与口令等。而现在，企业需要保护三类敏感数据，包括企业敏感数据、个人隐私数据和国家敏感数据。后两者受法规的监管与保护：欧盟 GDPR、美国 CCPA 的法规中明确表示公民的个人隐私数据受保护；我国法规监管对象不仅包括个人隐私数据，还包括各类国家敏感数据，法规上也称为“重要数据”，比如未公开的政府信息，大面积人口、基因健康、地理、矿产资源等。法规定义的个人数据 / 个人信息不是传统意义上的身份证号、手机号、地址等个人基本信息，还包括设备的 IP 地址、MAC 地址、Cookie 信息，范围非常宽广，这给企业的敏感数据识别和保护构成巨大的挑战。

（2）涵盖的应用场景更加多种多样：原先企业以重要资产的数据为数据安全保护对象，主要面向的是数据库和机密文档等环境。而现在，法规监管的基本单位是数据，而不仅是数据库和文档的数据，还包括大数据平台、文档、云、终端，甚至包括发展中的 5G、区块链等新型环境的符合法规定义的个人隐私与重要数据。

图 1 数据安全涵盖各类环境

（3）涵盖数据的全生命周期：原先企业数据安全主要面向数据传输和数据存储过程。现在，数据安全覆盖数据的全生命周期的各个环节，包括数据采集、传输、存储、处理、交换和销毁。比如在数据采集时，必须遵循最小可用、征得用户同意等合规性原则，进行相应数据安全建设。

图 2 数据安全涵盖全生命周期

在合规视角下，数据安全的需求和驱动力发生了根本性的改变，数据安全保护的数据对象范畴变得更大，数据安全覆盖的应用场景将变得更加多样，数据安全需求将覆盖数据的全生命周期。为了更好地应对合规性带来的挑战，一方面企业需要从传统单点的数据安全建设转变成体系化的数据安全建设，另一方面为了应对三类重要合规场景带新来的挑战，包括用户隐私数据安全合规、企业内部数据安全治理、企业间数据安全共享与计算，亟须引入一些前沿的数据安全技术实现困境的破局。

参考文献

[1] https://gdpr.eu/privacy-notice/.

[2] https://oag.ca.gov/privacy/ccpa.

11月1日，绿盟科技**数据要素流通安全能力**正式发布。

今日热点

数据安全「3+2」

如何为数据流通的各类应用场景提供高度可靠、安全的能力支持？

3大数据安全产品

API 安全更进一步

API 风险监测与审计系统（APISEC）聚焦网络中的 API 流量与敏感数据，发现 API 自身安全风险与 API 数据泄露风险，具有 API 资产识别准确率高、敏感数据类别覆盖广、异常行为发现模型丰富的技术优势。

要监测也要防护

数据安全网关（DSG）提供数据分类分级、脱敏、审计和访问控制等能力的（1+N）一站式数据安全防护产品，解决数据流转过程中的数据调用和应用安全。

1个 Agent，对终端数据威胁说“不”

下一代电子文档加密系统（CDG 融合版）融合文档加密、数据防泄露和终端桌面管理等功能，以终端数据资产为核心，全方位保障用户终端数据安全。

2大数据安全解决方案

更智能 更高效 更全面 更精准

实战化防泄露解决方案从攻击者视角识别敏感数据内容和行为，利用威胁情报信息，实时判断高危访问源的风险行为，让数据防泄露进入常态化防御状态。

随意 无序 → 可信 可控

可信数据空间解决方案通过开发与生产环境隔离、租户隔离，以及可信执行环境、机密计算技术，为数据授权运营提供安全可信的数据加工利用环境，以此实现数据高效合规流通使用。

THE EXPERT
BEHIND GIANTS
巨人背后的专家

多年以来，绿盟科技致力于安全攻防的研究，为政府、金融、运营商、能源、交通、科教文卫等行业用户和各类型企业用户，提供具有核心竞争力的安全产品及解决方案，帮助客户实现业务的安全顺畅运行。在这些巨人的后面，他们是备受信赖的专家。

客户支持热线：400-818-6868

NSFOCUS 绿盟科技