安全+
SECURITY
技术版 与安全人士分享技术心得 Share technique experience with security professionals
2024/07 总第 061

★ 本期焦点

构建安全壁垒：Kubernetes攻击与防御技术综述

工业控制系统网络安全防御体系思考

网络安全政策导读（2024年1—6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号益泰大厦三层
邮编：100089
电话：(010)6843 8880-5462
传真：(010)6872 8708
网址：www.nsfocus.com

2024/07 总第 061

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

作为全球规模最大的网络安全盛会之一，2024年RSA大会于5月9日落下帷幕。在大模型日新月异发展的今天，人工智能再次站在科技舞台的中央，创新、商机与威胁成为网络安全的关键词。

本期《安全+》将继续立足网络安全发展，从前沿技术发展、安全理念应用、政策解读等视角出发，在AI重新定义网络安全的大背景下，洞见网络安全背后的技术、思考与生态，分享网络安全洞察。

数字化浪潮下，AI能力已成为当下企业组建安全防护体系过程中不可或缺的一项技术。与过去相比，安全产品的能力和效率等方面均面临着更大的挑战。企业需要利用AI等技术，有效提升自身的安全防护等级。

2024年政府工作报告提出，大力推进现代化产业体系建设，加快发展新质生产力。新质生产力的高质量发展离不开高水平网络安全，只有实现安全产业的整体升级，才能支撑新质生产力的行稳致远。在此背景下，绿盟科技积极拥抱网络安全新质生产力，深化大数据、人工智能等在安全领域的研发应用，以创新驱动发展，以技术变革为路径，为锻造高水平网络安全能力持续发力。

叶晓虎

构建安全壁垒：Kubernetes 攻击与防御技术综述

绿盟科技 创新研究院 浦明

摘要：2024 年 RSA 大会上，来自 Google 的安全工程师 Lenin Alevski 和 Semgrep[1] 安全研究员 Max vonBlankenburg 的议题 Kubernetes Security:Attacking And Defending Modern Infrastructure[3] 从 Kubernetes 安全矩阵的各个阶段出发，向观众介绍了容器和容器编排工具的基本概念和组件功能，并详细介绍了 Kubernetes 的威胁模型，以及由 OWASP 提出的 OWASP Kubernetes Risks Top 10[2] 风险清单。 本文从 Kubernetes 威胁矩阵（ATT&CK）角度出发，阐述了针对 Kubernetes 的主要攻击与防御技术，同时对 Kubernetes 的未来演进趋势进行了介绍。

关键词：Kubernetes 安全 RSA 2024 容器安全 ATT&CK

1. 背景信息

云原生可称为云计算的下半场，近年兴起的容器和编排技术凭借其弹性敏捷的特性和活跃强大的社区支持，成为云原生生态的重要支撑技术。容器化部署形态也在改变云端应用的设计、开发、部署和运行，从而重构云上业务模式。

容器和容器编排系统的安全风险将直接影响整个云原生系统的安全性。从 IT 基础设施的视角看，云原生系统底层是容器，其基于操作系统虚拟化技术，跟其他的虚拟化云计算平台一样，存在逃逸和云内横向移动的风险；上层是以微服务为中心的容器编排、服务网格、无服务器计算（Serverless Computing）等系统，其 API、业务存在被攻击的风险。

此外，从 DevOps 的视角看，云原生系统所包含的软件供应链（如第三方软件库、容器镜像等、第三方厂商非授权发布软件仓库等）存在被投毒或恶意攻击的风险；整个开发环节，如 CI/CD，也存在被攻击的风险。

2.Kubernetes 威胁模型

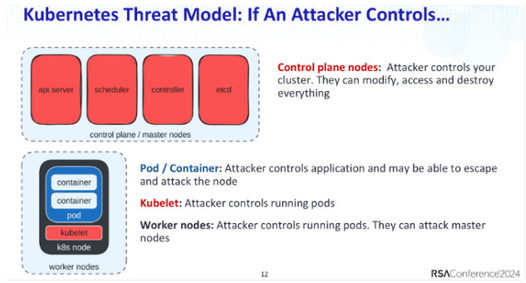


图 1 Kubernetes 威胁模型示意

Max vonBlankenburg 和 Lenin Alevski 提出了关于微软发布的 Kubernetes 的威胁模型，着重从攻击者的角度出发，分析了控制平面（Control Plane）和工作节点（Worker Nodes）两个攻击源的可能性。

如果攻击者控制了控制平面，他们可以对整个集群的主节点、工作节点以及其中运行的 Pod 和容器进行任意的访问、修改和破坏。

而如果攻击者控制了工作节点，则有以下两种可能的攻击手段：

1. 利用 Kubelet 组件控制运行的业务容器，对其进行增删改查。

2. 利用对容器中应用程序的已知漏洞或目录挂载进行利用从而达到容器逃逸至宿主机的目的。容器逃逸比较知名的 CVE 有 2016 年曝出的脏牛漏洞（CVE-2016-5195）及 2019 年曝出的 runC 漏洞（CVE-2019-5736）。关于容器逃逸漏洞，绿盟科技研究通讯曾发布过相关文章《云原生攻防研究——容器逃逸技术概览》《容器逃逸成真：从 CTF 解题到 CVE-2019-5736 漏洞挖掘分析》，感兴趣的读者可以阅读。

3.Kubernetes 威胁矩阵

在他们的演讲中，安全研究人员首先介绍了微软在 2020 年提出的 Kubernetes 威胁矩阵，然后与最近一次微软在 2021 年提出的新版本进行了比较，使得观众能够更清楚地了解 Kubernetes 安全领域的发展和演变。

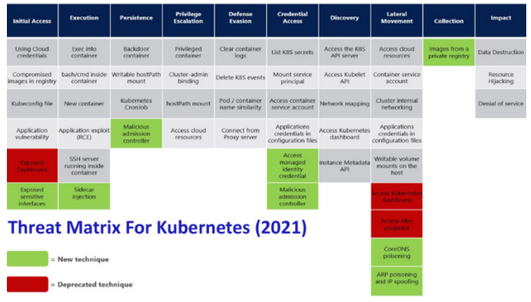


图 2 微软在 2021 年提出的 Kubernetes 威胁矩阵

可以看出，相较于旧版本，新版本有以下变化：

- **初始访问矩阵**

废弃了“暴露的 Kubernetes Dashboard”攻击面。实际上，从 Kubernetes Dashboard 的历史版本迭代中可以看出，在 v1.10.1 之前，无须配置登录即可访问，仅需在 Web 登录界面点击“跳过”按钮。但在 v1.10.1 版本后，开发团队增加了显式配置功能，需要在相应部署的 yaml 文件中指定 --enable-skip-login 参数才能开启未授权访问。而当前的 Kubernetes Dashboard 已经延伸至 v7.4.0 版本，旧版本的使用率已经大幅降低。

另外，新增了“暴露的敏感信息接口”这一攻击面。例如，许多私有化部署的容器编排平台，如 Openshift、Rancher、VMware Tanzu 等，都提供了敏感信息接口。攻击者可以利用这些接口的不安全配置对平台进行未授权访问。

- **执行矩阵**

新增了“边车容器注入”的攻击面。最初，边车容器实际上是作为辅助容器与主业务容器在同一 Pod 中运行，以此来增强或扩展主应用容器的功能，而无须直接修改主应用代码。然而，近年来随着服务网格的流行（例如 Envoy Sidecar 模式），边车容器也常被作为一种攻击技术。主要目的是提升攻击资产的隐蔽性，使得攻击更加难以检测和防御。

- **获取凭证矩阵**

新增了“恶意准入控制器”和“访问管理身份凭证”攻击面。特别需要注意的是“恶意准入控制器”，它实际上是一段代码，在用户请求通过认证授权之后，Kubernetes 资源对象持久化之前进行拦截。集群管理员可以通过在 kube-apiserver 配置文件中指定“--enable-admission-plugins”参数项的值来启用准入控制器。Kubernetes 包含许多准入控制器，其中常见的有 MutatingAdmissionWebhook 和 ValidatingAdmissionWebhook 准入控制器。它们分别用于拦截并修改 Kubernetes API Server 请求的对象以及对其对象格式进行校验。近年来，利用准入控制器修改 YAML 以向业务 Pod 中添加恶意容器的攻击面变得尤为突出。

- **横向移动矩阵**

废弃了“Kubernetes Dashboard 的访问”和“访问 Tiller 端点”的攻击面。访问 Tiller 端点的攻击面与前文提及的 Kubernetes Dashboard 类似，都是因为 Kubernetes 社区和相关厂商对 Tiller 的安全性问题进行了持续改进，可以通过更新和加固 Tiller 组件来消除这些攻击面。

新增了“CoreDNS 投毒”和“ARP 投毒”攻击面。CoreDNS 是 Kubernetes 中使用的主要 DNS 服务。攻击者可以通过修改位于 kube-system 命名空间的 ConfigMap 中的 corefile 文件来更改集群 DNS 的行为，从而对其进行投毒，并获取其他服务的网络身份。

- **收集矩阵**

新增了“获取私有镜像”攻击面，公有云托管的 Kubernetes 集群服务中，如果攻击者可以访问集群，某些情况下则可能对相应容器镜像仓库进行访问，从而对镜像进行投毒操作。

接下来，笔者将对本次议题的重点—— 常见的 Kubernetes 攻击技术进行介绍。

3.1 初始访问矩阵

常见的攻击技术包括利用云凭证、暴露的 Kubeconfig、受攻击的镜像仓库 / 镜像、利用已知漏洞、暴露的敏感接口等。

(a) 利用云凭证可以访问公有云提供的 Kubernetes 服务。目前，由于安全意识不足，开发者常将云凭证硬编码到系统业务源码中，导致这些敏感信息在公网泄露。攻击者可以利用泄露的云凭证访问 Kubernetes 资产，并进行篡改、删除等恶意操作。

(b) kubeconfig 与暴露的云凭证类似，只是攻击者可以通过 kubeconfig 在网络可达的情况下获取 Kubernetes 控制面的权限，从而对集群资源进行恶意操作。

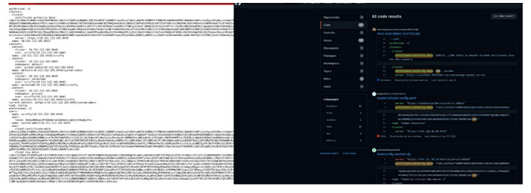


图 3 泄露的 kubeconfig 信息

(c) 利用受攻击的镜像仓库，上传恶意镜像并进行投毒，引发供应链攻击。

(d) 暴露的敏感接口以及已知漏洞利用同样会导致未经授权的访问和容器逃逸，从而增加横向移动的风险。

3.2 执行矩阵

本次议题中，Max vonBlankenburg 和 Lenin Alevski 安全研究员介绍了执行矩阵的一些攻击技术，主要分为三个场景：

场景一：利用用户凭证、kubeconfig、应用 RCE 漏洞进行资产探测。

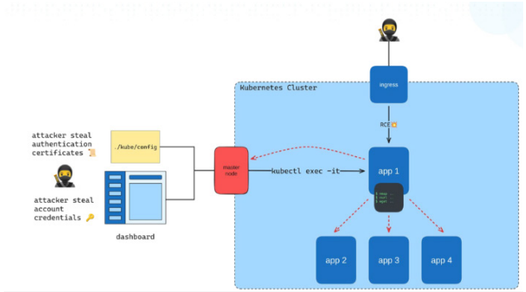


图 4 执行矩阵 - 攻击场景 1

如图 4 所示，攻击者主要通过用户凭证、kubeconfig 等信息获取到 Kubernetes 主节点的控制权，并通过 kubectl 访问业务 Pod 执行资产探测操作，或是通过业务 Pod 的已知 RCE 漏洞在 Pod 内部进行命令执行操作。

场景二：利用用户凭证、kubeconfig 创建恶意容器进行挖矿。

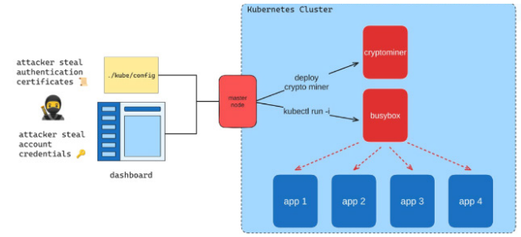


图 5 执行矩阵 - 攻击场景 2

如图 5 所示，攻击者主要通过用户凭证、kubeconfig 等信息获取到 Kubernetes 主节点的控制权，并通过 kubectl 创建 crypto miner Pod 执行恶意挖矿操作。

场景三：利用用户凭证、kubeconfig 在现有业务 Pod 上注入边车容器进行恶意操作。

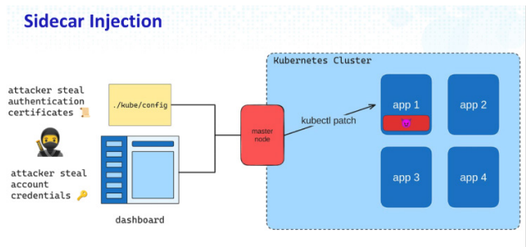


图 6 执行矩阵 - 攻击场景 3

如图 6 所示，攻击者主要通过用户凭证、kubeconfig 等信息获取到 Kubernetes 主节点的控制权，并通过 kubectl patch 在现有 Pod 基础上注入 Sidecar 容器执行恶意操作。

3.3 持久化矩阵

Max vonBlankenburg 和 Lenin Alevski 安全研究员在持久化矩阵介绍了两种攻击场景。

场景一：利用用户凭证、kubeconfig、应用 RCE 漏洞在现有业务 Pod 中挂载 Hostpath 写入操作。

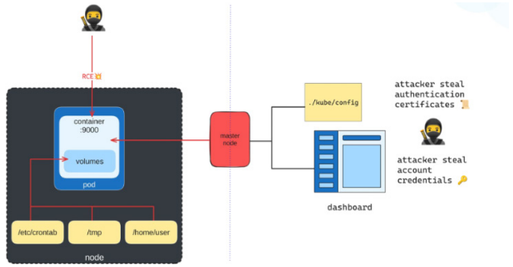


图 7 持久化矩阵 - 攻击场景 1

如图 7 所示，攻击者主要通过用户凭证、kubeconfig 等信息获取到 Kubernetes 主节点的控制权，并通过 kubectl 修改业务 Pod 的 yaml 配置文件，将容器内的可写目录以 Hostpath 形式挂载至宿主机，从而进行持久化操作。

场景二：通过 Malicious 准入控制器在现有业务 Pod 中注入恶意边车容器。

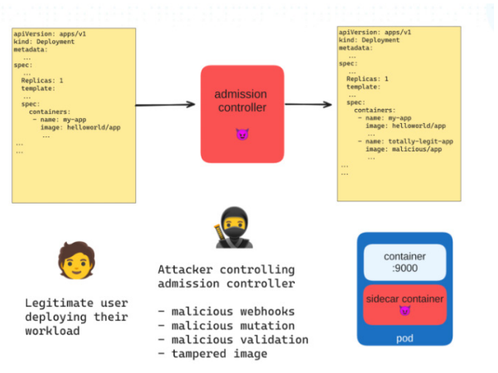


图 8 持久化矩阵 - 攻击场景 2

如图 8 所示，攻击者可通过前述的凭证和 kubeconfig 配置文件获取到 Kubernetes 主节点控制权，通过创建 Admission Controller 拦截正常用户部署的工作负载请求，并在请求所属业务 yaml 配置文件中注入边车容器配置，从而进行持久化操作。

3.4 权限提升矩阵

场景一：部署特权容器导致权限提升

如我们所知，一旦 Pod 部署时指定 Privleged 配置为 True, 则该 Pod 与宿主机共享 PID、Network 等命名空间，因此攻击者可通过共享命名空间的方式提权达到效果。

场景二：绑定集群角色权限导致权限提升

若Pod所属服务账户(Service Account)绑定了Kubernetes的ClusterRole 权限，同时该 Pod 存在 RCE 漏洞，则攻击者可以通过服务账户的 Token 操作集群资源，进而达到权限提升的效果。

3.5 防御绕过矩阵

清理容器日志、删除 Kubernetes 事件:避免日志审计进行溯源;

利用 Pod/ 容器名的相似性：提升攻击资产的隐秘性；

使用代理服务连接集群：避免对访问记录进行溯源。

3.6 获取凭证矩阵

通过 kubectl 列出 Kubernetes 的 Secrets 资源信息；

通过 kubectl 访问 Pod 服务账户（Service Account）；

通过部署恶意准入控制器获取凭证信息。

3.7 发现矩阵

Kubernetes 的核心组件 API Server 和 Kubelet 可用于发现集群中的元数据信息，如 Deployment、Pod、Service 等业务资源，同时也可以通过 Kubernetes Dashboard 或 Instance Metadata API 发现更多的集群资源信息。

3.8 横向移动矩阵

利用应用漏洞进行容器逃逸并访问集群中的其他资源，达到横向移动的效果；

利用服务账户进行横向移动；

利用集群内部的网络进行横向移动；

利用集群业务应用中的凭证信息进行横向移动；

利用 CoreDNS 投毒进行横向移动。

3.9 收集矩阵

演讲者主要提到通过访问私有镜像仓库，进行镜像下载，并通过镜像投毒的方式引发供应链攻击，上文已有说明，此处不再赘述。

3.10 危害矩阵

攻击者可能删除 PV（持久卷）、日志、事件等资源；

攻击者可能进行 Kubernetes 资源劫持；

攻击者可能利用上述的边车注入或创建容器的方式部署加密货币挖矿程序，进行种子下载等恶意操作；

攻击者在控制 Kubernetes 管理平面后可对 Pod 进行恶意操作，并部署新的恶意服务。

4.Kubernetes 内部安全机制

- **API Server认证授权**

API Server 实现了 Kubernetes 资源增删改查的接口，因而在用户对资源进行操作之前，需要对用户进行相应的认证操作，Kubernetes 支持多种认证类型，例如静态令牌文件、X.509 客户端证书、服务账号令牌、基于 OAuth 2.0 的 OpenID Connect（OIDC）令牌、Webhook 令牌身份等。

- **API Server授权（Authorization）**

Kubernetes 包含四类授权模式，分别为节点（Node）授权、基于属性的访问控制（Attribute-based access control,ABAC）、基于角色的访问控制（Role-based access control,RBAC)、基于钩子（Webhook）方式的授权，目前业界使用，RBAC 机制较多。

- **准入控制器（Admission Controller）**

Kubernetes 包含许多准入控制器，其中有两个特殊的准入控制器：

（1）MutatingAdmissionWebhook

变更准入控制器，可以拦截并修改 Kubernetes API Server 请求的对象。

（2）ValidatingAdmissionWebhook

验证准入控制器，可以对 Kubernetes API Server 请求对象的格式进行校验，但无法修改对象。

集群管理员可通过修改 kube-apiserver 配置文件启动以上两个准入控制器：

--enable-admission-plugins=NodeRestriction，MutatingAdmissionWebhook,ValidatingAdmissionWebhook

准入控制过程主要分为两个阶段，第一阶段运行变更准入控制器，第二阶段运行验证准入控制器，需要注意的是，Kubernetes 的某些准入控制器既是变更准入控制器也是验证准入控制器。如果第一阶段的任何准入控制器拒绝了请求，则整个请求被拒绝，并同时会向终端用户返回一个错误。

- **Secret**

Kubernetes 使用 Secret 对象来保存敏感信息，例如密码、令牌和 SSH 密钥。相比于将敏感信息放入 Pod 定义的 yaml 文件或容器镜像中，使用 Secret 方式更为安全灵活，该方式也是目前开发者常使用的密钥管理方式。在传递至容器的过程中，主要有将 Secret 构建至容器镜像中，通过 Kubernetes 环境变量，挂载宿主机文件系统三种方式。

- **网络策略（Network Policy）**

随着应用在云原生环境中的逐步落地，应用上云后带来了诸多网络层面的问题，例如应用间的网络调用需求大规模增长，应用间的流量控制变得尤其复杂，面对这些问题，Kubernetes 在 1.3 版本引入了网络策略，其主要用于在网络三四层提供流量控制能力，网络策略以应用为中心，允许用户设置 Pod 与网络中各类实体间的通信。

网络策略需要通过网络插件来实现，由于某些网络插件不支持网络策略，例如 Flannel，所以策略即使下发成功也不会生效。支持网络策略的网络插件包括但不限于 Calico、Cillum、Weave 等。

默认情况下，Pod 间是非隔离的，接受任何来源的流量。当为某一命名空间下的 Pod 下发网络策略时，该 Pod 会拒绝该网络策略所不允许的连接，其他命名空间的 Pod 继续接收流量。此外，网络策略通常不会冲突，是累加的并且最终取并集，不会影响策略结果。

5.Kubernetes 未来演进趋势

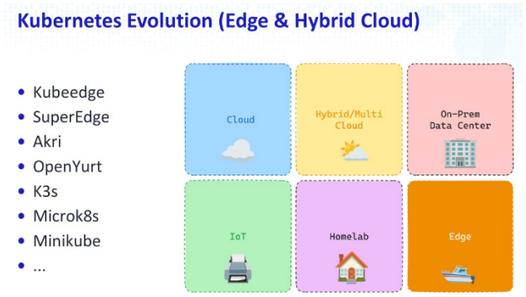


图 9 Kubernetes 未来演进趋势示意

在本次 RSA 演讲中，Max vonBlankenburg 和 Lenin Alevski 谈到 Kubernetes 的演进趋势，认为 Kubernetes 将会以更多的形态（K3s、Minikube、Kubeedge）部署在边缘云和混合云场景中，笔者也对此表示认同，未来的云原生技术将赋能于各类新型基础设施，如 5G、边缘计算、工业互联网等，云原生安全将根据不同场景的特点、需求和约束条件，演化出多种技术发展路线，最终形成相应的新基建安全防护方案。此外，随着企业上云、SDWAN 兴起，安全访问服务边缘（Secure Access Service Edge，SASE）技术将会把各类安全技术与网络、云基础设施融合，提供融合私有云、公有云、多云、混合云和传统 IT 环境等复杂场景下统一的端到端安全连接，这种环境变化，将催生新形态的安全能力和安全交付模式。

6. 总结

从今年的 RSA 议题和创新沙盒入围的安全初创公司可以看出，云安全仍然是热门话题之一。云原生概念自被提出以来，企业业务云原生化改造在国内外各行业都已得到大规模实施。安全伴随业务，随之而来的便是云原生安全，特别是容器安全和 Kubernetes 安全。Gartner 近年也提出了 CNAPP 的概念，这预示着云原生防护体系将会不断完善。本次议题中 Max vonBlankenburg 和 Lenin Alevski 全面阐述了 Kubernetes 的常见攻击手法，并提出了相应的安全建议，为企业在云原生安全实际落地过程中提供了有价值的参考。

参考文献

[1] https://semgrep.dev/about .

[2] https://owasp.org/www-project-kubernetes-top-ten/ .

[3] RSAC 2024 Kubernetes Security Attacking and Defending Modern Infrastructure.pdf.

UEFI Bootkits，当安全启动不再“安全”

绿盟科技 创新研究院 高翔

摘要：2024 年 RSA 大会上，研究员 Martin Smolar 分享了有关 UEFI 安全启动和 Bootkits 的议题。UEFI 安全启动（Secure Boot）旨在确保计算机 UEFI 固件启动的代码可信，从而保护系统，防止恶意代码在操作系统加载前的启动过程中被加载和执行。此前的系列文章已介绍了安全启动的基本知识 [1][2]，而本文将着重讨论安全启动涉及的密钥、潜在的安全漏洞以及加固方法。

1. 什么是 UEFI 安全启动?

几十年来，个人电脑一直受到病毒、蠕虫和其他恶意软件的困扰。最早的一些个人电脑病毒是以引导扇区病毒的形式传播的：它们以代码形式存在于软盘的引导扇区中，当用户使用受感染的 DOS 软盘启动计算机时，病毒就会从一台计算机传播到另一台计算机。虽然随着软盘重要性的降低和互联网连接的普及，其他病毒传播方式也逐渐显现，但系统启动前阶段（pre-boot）的恶意软件对黑客来说始终具有吸引力。通过在操作系统内核获得计算机控制权之前执行，恶意软件可以“隐藏”起来，而一旦操作系统获得控制权，恶意软件就无法“隐藏”起来。因此 pre-boot 阶段的恶意软件很难被检测到（除非重启到未受影响的应急系统里）。

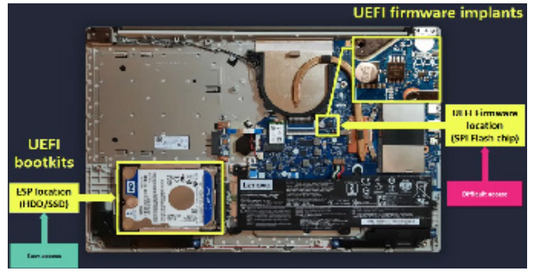


图 1 UEFI 固件的位置

BIOS/UEFI 几乎无法防止启动前的恶意软件；在 BIOS/UEFI 启动过程中，系统默认信任 boot loader 执行的任何程序。直到 2012 年年底，大多数的 EFI 实现也是如此。传统的系统启动分为以下几个过程:开启电源——UEFI 固件——硬盘中的操作系统(boot loader、内核)。

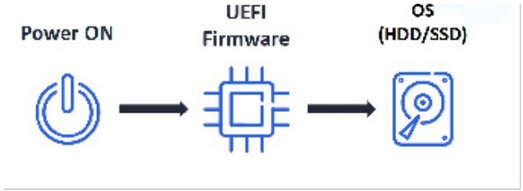


图 2 UEFI 启动过程

在系统启动过程中，UEFI 固件会根据 Boot variables 决定启动顺序，并且执行磁盘 ESP 分区中的 UEFI 应用。如图 3 所示，攻击者可能修改 Boot order 以及 ESP 分区中的二进制文件。

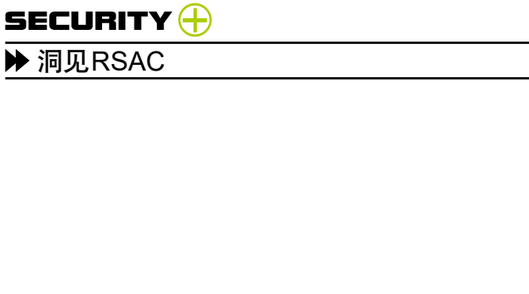


图 3UEFI 潜在攻击点

安全启动旨在为 pre-boot 过程添加一层保护。开启安全启动后，固件会检查执行的任何 EFI 程序是否存在签名。如果签名不存在，与计算机 NVRAM 中的密钥不一致或被列入 NVRAM 的黑名单，固件就会拒绝执行该程序。一个可信的 EFI boot loader 必须以安全的方式继续引导，最终实现一个安全的操作系统。

安全启动中的重要概念是允许签名数据库 (db) 和禁止签名数据库 (dbx)。允许签名数据库存储机器固件允许加载的受信任 boot loader 和 EFI 应用程序的哈希值和证书。禁止签名数据库存储已撤销、受损和不可信任的哈希值和证书。任何使用禁止 dbx 密钥加载签名代码的尝试，或在哈希值与禁止 dbx 条目匹配的情况下，都会导致启动失败。对 db 和 dbx 的签名，需要用到密钥注册密钥数据库 (KEK) 和平台密钥数据库 (PK)。

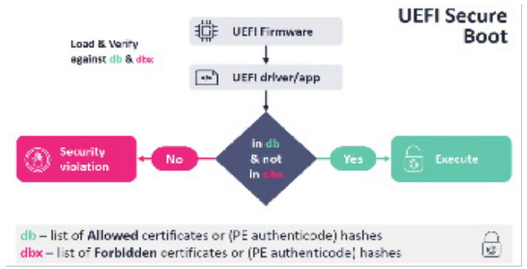


图 4 db 与 dbx

这些密钥用于签署启动加载程序、驱动程序固件运行的其他软件。目前销售的大多数商品 PC 都包含由微软控制的密钥。

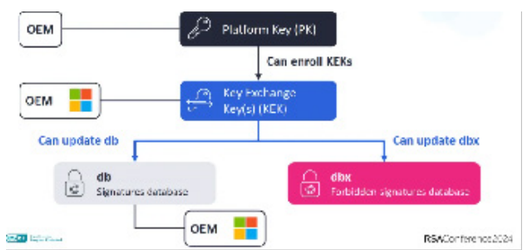


图 5 安全启动涉及的密钥

2. 开启了安全启动的系统就绝对安全吗?

只要开启了安全启动的系统就绝对安全吗? 我们总结了几类涉及 UEFI 安全启动绕过的漏洞。

内存相关漏洞：2022 年 1 月，研究者发现了名为“baton drop”的安全启动漏洞 (CVE-2022-21894) [3]。Windows 启动应用程序允许通过设置“truncatememory”移除包含序列化数据相关的内存块，从而绕过安全启动。Truncatememory BCD 元素将从内存映射中删除指定物理地址以上的所有内存内容。在从内存读取序列化的安全启动策略之前，攻击者会在初始化过程中对每个启动应用程序执行此操作。尽管微软很快修复了此漏洞，但由于受影响的 UEFI 二进制没有被撤销，“baton drop”攻击在很长的一段时间内仍然可被利用，产生了相关联的漏洞 [4]。直到 2023 年 5 月，微软才给出处理建议 [5]。

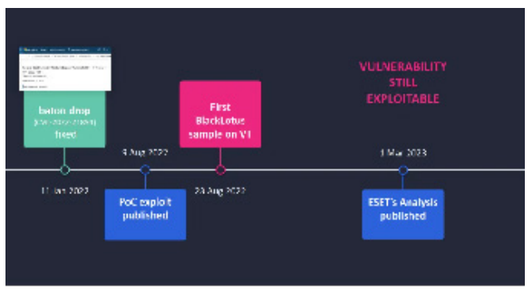


图 6 微软修复过程

特权命令、撤销列表相关漏洞：在 2020 年研究者发现，攻击者可以通过 insmod 加载一个未被签名的 grub module 来绕过安全启动 [6][7]。微软花费数月才彻底地将此漏洞移除。类似地，UEFI shell 中也存在大量的敏感命令，在开启安全启动的设备中，这些命令必须被禁止 [8][9]。除此之外，一些第三方的 PE 也被爆出存在安全风险，他们从硬编码的地址，加载运行任意的 UEFI 二进制 [10]。

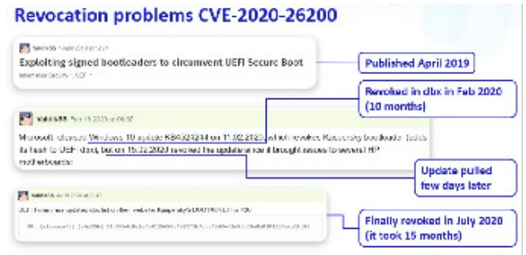


图 7：微软修复涉及撤销相关漏洞的过程

Debug 相关漏洞：ESET 的研究人员发现某些品牌笔记本的 UEFI 固件存在安全风险，攻击者可以直接从 OS 中创建 NVRAM 变量来关闭 UEFI 安全启动或者恢复出厂设置 [11]。

图 8 UEFI 安全启动漏洞

3. 如何控制安全启动的密钥来预防 Bootkits？

传统的解决方案，例如“最新版本的系统更新、更新固件、开启安全启动、安装安全防护软件等”往往是不够的。这是由于 Bootkits 往往需要花费数月的时间去彻底移除，移除可能不彻底，签名过程不透明。因此我们需要尽早地介入，例如可以阻断任何试图修改 ESP 分区文件的行为，可以通过白名单的方式来实现它。除此以外，还可以自定义安全启动：完全定制化和部分定制化。

完全定制化的安全启动将会移除所有安全启动的密钥，并且只使用用户的密钥，所有启动的组件都由客户的密钥签名。完全定制化的安全启动对用户而言，往往很难维护。

而部分定制化的安全启动则允许用户将自己的密钥追加进 PK、KEK 数据库中，当发现安全漏洞时，用户可以提前一步，将恶意程序的 hash 写入 dbx 中，避免了微软 /OEM 厂商的漫长的响应。而对于只用 Linux 的用户而言，还可以将 Windows UEFI CA 从 db 数据库中移除，进而减少攻击面。

从上述的自定义安全启动可以看出，其实质是重新控制安全启动的密钥。笔者总结了如下益处：

从默认密钥中移除风险

理论上，安全启动应能阻止恶意软件运行。但攻击者总是有可能诱骗微软签署恶意软件，或者签署的软件可能存在漏洞，比如 2020 年发现的 Boot Hole 漏洞。如果使用默认密钥的 Shim，并且 dbx 没有进行对应的更新，计算机仍会受到这些威胁的攻击。

从发行版密钥中移除风险

与前述情况类似，发行版密钥也有可能被泄露，在这种情况下，攻击者可能会分发使用泄露密钥签名的恶意软件。

消除对 MOK 的需求

Shim 和 PreLoader 工具都依赖于机器所有者密钥 (MOK)，MOK 与安全启动密钥类似，但更容易安装。由于更容易安装，它们更容易被社会工程或其他手段滥用。尤其是在管理一系列由其他人使用的台式电脑时，取消 MOK 可以增加安全性。

测试和开发

如果想开发自己的启动管理器，可能需要在模拟环境中测试软件的签名版本。不过，使用微软安全启动密钥签署二进制文件的过程烦琐而耗时，因此可能需要用自己的密钥以便自己签署二进制文件。当软件按照预期运行时，就可以将它发送给微软进行签名了。

解决启动问题

在双系统的场景中，某些电脑默认只能启动 Windows，尽管可以暂时启动到 Linux，使用 efibootmgr 将 Linux 设置为默认启动加载器，但随后发现自己又启动回 Windows，因为固件一直将 Windows 作为默认设置。如果 Linux 启动项仍然存在但被“降级”，那么设置自己的启动密钥就可以解决这一问题，具体方法是从常规安全启动列表中删除微软的密钥，然后将其添加到 MOK 列表中。

4. 总结

前几年对于 Linux 用户来说，安全启动介于“无所谓”和“很麻烦”之间。虽然安全启动有可能提高安全性，但相关安全风险没有受到大家的重视。因此，虽然安全启动在理论上有好处，但对于只使用 Linux 的计算机来说，安全启动是否有实际好处还不清楚。随着 Bootkits 的出现，人们愈加重视安全启动以及对应的漏洞所造成的安全风险。厂商无法及时地将恶意软件加到撤销列表中，就导致了许多在野漏洞的出现。签署你的 boot loader 以使用自己的密钥可以提供最大的安全性和灵活性。绿盟科技会持续关注安全启动的进展，欢迎感兴趣的读者持续关注。

参考文献

[1] https://mp.weixin.qq.com/s/wucSVNYYeg5d5fQ1l1cnAQ.

[2] https://mp.weixin.qq.com/s/showAKatT3TsN11aWRD9GQ.

[3] https://nvd.nist.gov/vuln/detail/CVE-2022-21894.

[4] https://nvd.nist.gov/vuln/detail/CVE-2023-24932.

[5] KB5025885: How to manage the Windows Boot Manager revocations for Secure Boot changes associated with CVE-2023-24932 - Microsoft Support.

[6] https://nvd.nist.gov/vuln/detail/CVE-2020-7205.

[7] https://nvd.nist.gov/vuln/detail/CVE-2020-26200.

[8] https://nvd.nist.gov/vuln/detail/CVE-2022-343010.

[9] https://nvd.nist.gov/vuln/detail/CVE-2022-34303.

[10] https://nvd.nist.gov/vuln/detail/CVE-2022-34302.

[11] https://nvd.nist.gov/vuln/detail/CVE-2022-3431.

供应链与AI安全挑战

绿盟科技 创新研究院 卜天

摘要2024 年 RSA 大会上，来自泰雷兹(THALES)的软件安全技术总监 Viswanath Chirravuri 分享了关于供应链安全和 AI 安全的议题。泰雷兹是一家业务遍及全球的国际企业，所服务的五大业务市场对各国社会至关重要，包括航空、航天、地面交通、防务与安全以及数字身份与安全。网络攻击更迭变化，已经从攻击计算机的硬件、网络、软件，转变成攻击供应商所提供的任何内容，包括数据、服务、代码、组件、配置、二进制、进程、人等。自 2022 年年底大模型技术迅速发展，如何理解 AI 与安全的关系成为一个重要的课题，本文将从供应链和 AI 安全两个方面分别进行解读。

1. 供应链攻击与分类法

供应链攻击 (Supply chain attacks) 多年来一直是一个安全问题，自 2020 年以来全球频发供应链攻击事件。2020 年 12 月，SolarWinds 公司的 NMS 产品因受第三方应用漏洞的影响而被盗，进而导致泄露的软件被客户直接下载。2020 年 7 月，攻击者通过有效凭证访问到 Ledger 公司的电子商务数据库，利用盗用的数据进行网络钓鱼和勒索活动。2021 年 7 月，攻击者利用远程监控公司 Kaseya 的系统漏洞，向所有 VSA 设备发送远程更新并执行攻击者代码，最终对 VSA 系统管理的客户发起勒索。

随着安全意识的提升，企业和组织实施了更强大的安全保护策略，使得攻击者将攻击目标转向了供应商，并设法在系统停机、金钱损失和声誉损害等方面造成重大影响。供应链攻击利用了全球市场间相互连接的关系，当多个客户依赖于同一供应商时，针对该供应商的网络攻击风险就会被放大，有可能导致全国性甚至全球性的大规模影响。

从攻击特点上看，供应链攻击至少包含两个步骤，一次针对供应商，另一次针对攻击目标，攻击目标可以是最终客户或其他供应商。因此，要将攻击归类为供应链攻击，供应商和客户都必须是目标，通过前后攻击两个目标来达到访问资产的目的。

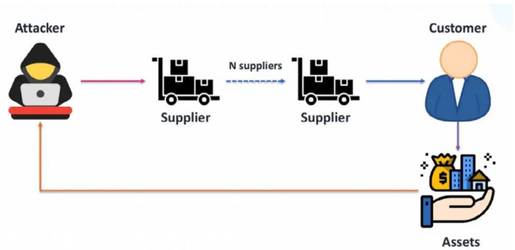


图 1 供应链攻击步骤

报告提出了一种分类法 [1] 来描述供应链攻击，这个分类法考虑了供应链的四个关键元素，包括：i. 对供应商使用的攻击技术；ii. 供应商被攻击的资产；iii. 对客户使用的攻击技术；iv. 客户被攻击的资产。这种分类可以帮助组织理解供应链攻击的各个部分，并将它们与其他类似的网络攻击进行比较，更重要的是将这些事件识别为供应链攻击。

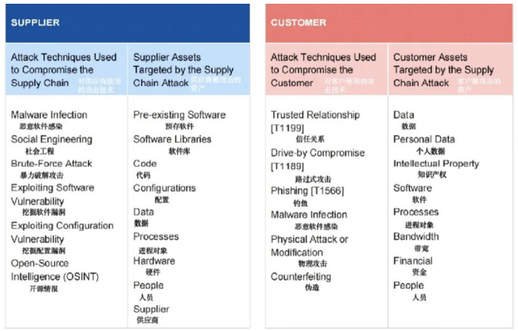


图 2 供应链攻击分类法

每种技术用于识别攻击是如何发生的，而不是攻击的目标是什么，任何供应链攻击中都可能使用了不止一种技术。具体而言，常用的供应商攻击技术包括恶意软件、社工、暴力破解、软解 \ 配置漏洞、物理攻击、开源情报、伪造等。

用户攻击技术包括信任关系攻击、路过式攻击、钓鱼、恶意软件、物理攻击、伪造等。

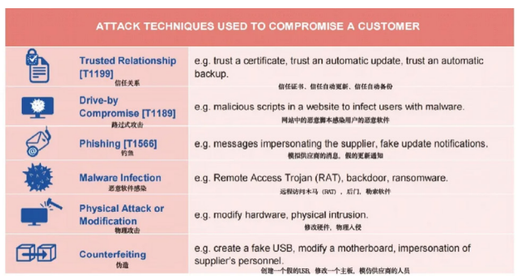


图 3 用户攻击技术

图 4 说明如何将分类法应用于真实情况，以用于识别和理解攻击的特定特征。Codecov 是一家提供代码覆盖和测试工具的软件的公司，该公司向 IBM 和惠普等公司提供工具。2021 年 4 月，Codecov 报告说，攻击者利用了一个在 Docker 图像创建方式的 bug，从 Docker 图像中获得了一些有效的凭证。通过这些凭证，他们破坏了 Codecov 客户使用的“上传 bash 脚本”。一旦客户下载并执行了这个脚本，攻击者就能够从 Codecov 的客户手中窃取数据，包括允许攻击者访问客户资源的敏感信息。

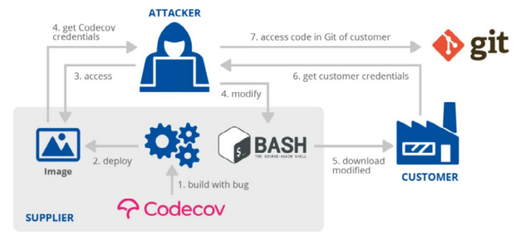


图 4Codecov 的供应链攻击事件

利用这些信息，可以识别出分类法中的四个元素。对供应商而言，本案例的攻击者通过“利用配置漏洞”来访问供应商，并在攻击中瞄准了供应商中的“代码”资产。对客户而言，攻击者利用了客户与供应商之间的“信任关系”，最终攻击了客户的源代码，因此目标是“软件”。

图 5 将 Codecov 事件应用到分类法

2. 软件供应链与标准

对于使用软件和可执行代码的用户，直接或间接地依赖于供应商提供的软件（包、库和模块）为他们带来极高的开发效率，但不透明的供应链和复杂的软件依赖关系也使用户的网络存在受到供应链攻击的安全威胁。报告列举了软件受供应链攻击的攻击面，共包含以下几个方面：

（1）应用程序源代码

被盗的代码签名证书或已签名的恶意应用程序、恶意软件插入 / 代码篡改。

（2）构建 CICS 系统

滥用 CI/CD 系统的默认行为、滥用网络挂钩来危害 CI/CD 系统、窃取凭据以在工件中注入恶意代码。

（3）源代码控制系统

滥用 Git 服务器配置错误、Git 存储库中不受信任的代码、向主分支中注入恶意代码。

（4）包管理器

前端组件（客户端）、后端组件（服务器端）、依赖性混淆。

（5）容器供应链

恶意的图像和脆弱的图像、不安全的容器注册表。

（6）云供应链

错误配置（公开的秘密、元数据服务等）。

（7）供应商的供应链

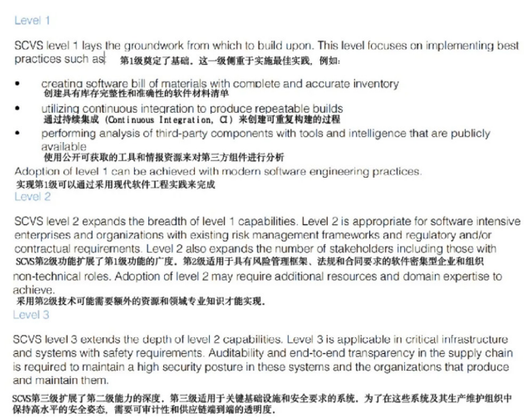
软件分发给用户前的供应链劫持。

为管理软件供应链中的风险，Viswanath 在会上提到了多种软件供应链的关键框架和标准，包括 OpenSSF sigstore、SLSA、OWASP Software Component Verification Standard (SCVS)、Google's software delivery shield (trusted OSS) 等，完整标准和链接参见文本参考文献处 [2]。

- OpenSSF sigstore
- SLSA
- Microsoft S2C2F
- NIST C-SCRM / SP 800-161
- NSA ESF (Enduring Security Framework)
- UK Supplier Assurance Framework
- MITRE System of Trust™ (SoT) Framework
- ISO/IEC 20243-1:2023 and ISO/IEC 27036
- SCS 9001 Supply Chain Security Standard
- OWASP Software Component Verification Standard (SCVS)
- Google's software delivery shield (trusted OSS)
- IETF SCITT

图 6 软件供应链的关键框架和标准

OSS 相关标准用于建立一个识别活动、控制和最佳实践的框架，以帮助识别和降低软件供应链中的风险。通过 OSS 标准可以减少易被利用的系统漏洞，并对衡量和改进软件供应链安全性非常重要。以 OWASP 提出的软件组件验证标准（SCVS）为例子，标准定义了三个验证级别，其中高级别包括低级别的控制范围。具体而言，SCVS L1 适用于最低标准，仅基本的分析形式就满足 L1 级别的软件。SCVS L2 需要额外的分析或详尽调查的中等敏感度的软件。SCVS L3 由于数据的敏感性或软件的使用，需要最高标准的要求。


Level 1

SCVS level 1 lays the groundwork from which to build upon. This level focuses on implementing best practices such as 第1级奠定了基础。这一级侧重于实施最佳实践。例如：

- creating software bill of materials with complete and accurate inventory
  创建具有库存完整性和准确性的软件材料清单
- utilizing continuous integration to produce repeatable builds
  通过持续集成（Continuous Integration，CI）来创建可重复构建的过程
- performing analysis of third-party components with tools and intelligence that are publicly available
  使用公开可获取的工具和情报资源来对第三方组件进行分析

Adoption of level 1 can be achieved with modern software engineering practices.
实现第1级可以通过采用现代软件工程实践来完成

Level 2

SCVS level 2 expands the breadth of level 1 capabilities. Level 2 is appropriate for software intensive enterprises and organizations with existing risk management frameworks and regulatory and/or contractual requirements. Level 2 also expands the number of stakeholders including those with SCS第2级功能扩展了第1级功能的广度。第2级适用于具有风险管理框架、法规和合同要求的软件密集型企业和组织non-technical roles. Adoption of level 2 may require additional resources and domain expertise to achieve.
采用第2级技术可能需要额外的资源和领域专业知识才能实现。

Level 3

SCVS level 3 extends the depth of level 2 capabilities. Level 3 is applicable in critical infrastructure and systems with safety requirements. Auditability and end-to-end transparency in the supply chain is required to maintain a high security posture in these systems and the organizations that produce and maintain them.
SCVS第三级扩展了第二级能力的深度。第三级适用于关键基础设施和安全要求的系统。为了在这些系统及其生产维护组织中保持高水平的安全姿态，需要可审计性和供应链端到端的透明度。

图 7 OSS 三层标准

然而，现有的 OSS 框架及标准更多地关注漏洞（可以被利用的弱点）而不是恶意点（故意设计的伤害）。漏洞库是指已知安全漏洞的库，这些漏洞可能被攻击者利用，其脆弱性是由编码错误、设计缺陷或其他问题造成。而恶意库是故意设计的，并通过恶意软件的分发来对用户造成伤害。对于漏洞库可使用 CVSS 评分来进行评价，而对于恶意库则缺少一套标准来进行规范。例如对一个 npm 恶意包，四个供应商使用了 4 个不同的符号进行标识。同时，大部分开源的 SCA 工具都不会报告恶意组件，这是一个需要关注的问题。

- cx-2021-b8833-be2146 (Checkmarx)
- SNYK-JS-RC-1911120 (Snyk)
- sonatype-2021-1696 (Sonatype)
- GHSA-g2q5-5433-rhrf (GitHub)

图 8 同一个恶意包的多种符号标号

3. AI 与网络安全

Viswanath 还介绍了人工智能（Artificial Intelligence, AI）与网络安全之间双向的互利关系。一方面，AI 能够极大地辅助网络安全领域开展入侵防御、威胁检测、漏洞识别、安全运营、取证等工作，这部分很多的网络安全公司都在做，绿盟科技也发布了网络安全大模型——风云卫。另一方面，随着大模型（Large Language Models,LLM）技术的快速发展，针对 AI 领域的新型攻击也层出不穷，如对抗性攻击防护、数据保护和隐私、供应链安全、模型安全等方面，都需要网络安全措施的引入和保护。

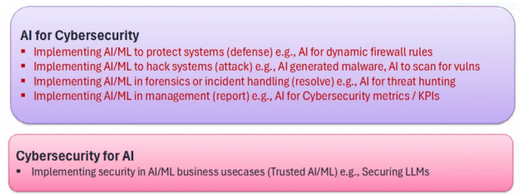


图 9 AI 与安全的双向关系

从 AI 自身安全的角度上讲有两种攻击目标，一种使用恶意的 AI 或机器学习(Machine Learning, ML）数据、模型攻击下游的公司、

系统、数据和用户；另一种使用对抗性的方法攻击 AI\ML 自身数据、模型和系统。在两方面交界处，也有使用恶意数据、模型作为渠道攻击其他 AI\ML 数据、模型的情况。

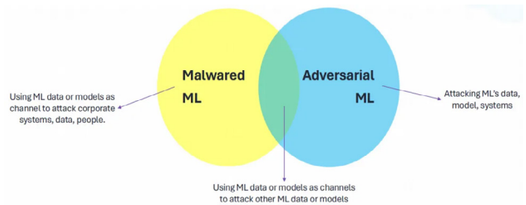


图 10 针对 AI 自身的两种攻击途径

结合 AI 自身安全和供应链安全而言，也有针对与 AI\ML 的供应链安全风险。现有做 AI\ML 相关研究所使用的数据、框架、预训练模型都由上游的供应方提供，如常用的开源框架 Tensorflow、pytorch、keras、scikit-learn 等，数据集如图像、NLP、语音等领域数据，公开可用的预训练模型可以从 Tensorflow Hub、Huggingface、OpenAI GPT、Torch Hub、Github 等社区获取。上述所提到的每一个节点都可能因受恶意代码的感染而将安全风险传递到下游模型，进而传递给万千大模型的使用者，最终造成不可估量的危害。

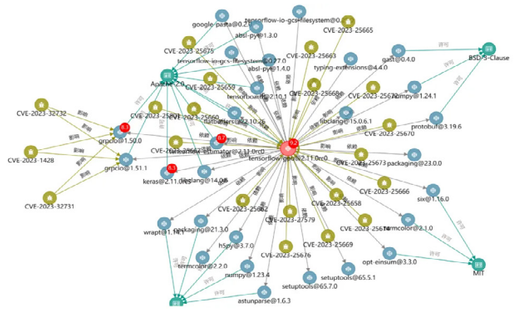


图 11 绿盟开源软件供应链平台展示的著名 tensorflow-gpu 组件 2.11.0rc0 版本受到 10 余 + 漏洞影响

在 2023 年 OWASP 提出的机器学习安全和大模型安全领域的 TOP 10 的攻击方式中，供应链攻击均榜上有名。

图 12 机器学习安全和大模型安全领域的 TOP 10 攻击

参考文献

[1] https://www.enisa.europa.eu/publications/threat-landscape-for-supply-chain-attacks.

[2] 软件供应链关键框架、标准和参考资料

OpenSSF sigstore

https://openssf.org/community/sigstore/.

SLSA

https://slsa.dev/.

Microsoft S2C2F

https://www.microsoft.com/en-us/securityengineering/opensource.

NIST C-SCRM / SP 800-161

https://csrc.nist.gov/projects/cyber-supply-chain-risk-management.

NSA ESF (Enduring Security Framework)

https://www.nsa.gov/About/Cybersecurity-Collaboration-Center/Enduring-Security-Framework/.

UK Supplier Assurance Framework

https://www.gov.uk/government/publications/government-supplier-assurance-framework.

MITRE System of Trust (SoT) Framework

https://sot.mitre.org/framework/system_of_trust.html

ISO/IEC 20243-1:2023 and ISO/IEC 27036.

https://www.iso.org/standard/82905.html.

SCS 9001 Supply Chain Security Standard

https://tiaonline.org/what-we-do/scs-9001-supply-chain-security-standard/.

OWASP Software Component Verification Standard (SCVS)

https://owasp.org/www-project-software-component-verification-standard/.

Google's software delivery shield (trusted OSS)

https://cloud.google.com/security/solutions/software-supply-chain-security?hl=zh-cn.

IETF SCITT

https://datatracker.ietf.org/wg/scitt/about/.

从“卷饼”中发现Kubernetes安全的钥匙

绿盟科技 创新研究院 张小勇

摘要：在企业 IT 基础设施建设中，Kubernetes 已成为企业部署和管理容器化应用的首选平台。然而，这也使得 Kubernetes 集群成为攻击者的目标。在今年的 RSA 大会上，Chipotle 的两位应用安全工程师 Raunaq Arora 和 Caleb Schwartz 分享了他们在保障 Kubernetes 安全方面的经验和最佳实践。本文将对该演讲进行深入解读，探讨 Kubernetes 环境中的安全挑战以及应对策略。

关键词：Kubernetes 零信任 K8S 微分段 服务网格

1. 背景介绍

Chipotle 是食品行业一家知名的公司，Raunaq Arora 和 Caleb Schwartz 是 Chipotle 的两位应用安全工程师。在保障其 IT 基础设施和应用安全的过程中，尤其是在 Kubernetes 环境的安全管理方面，他们积累了丰富的经验。

如图 1 所示，他们通过**对比 Burrito（卷饼）的制作过程和 Kubernetes 环境中的安全威胁**，以别具一格的方式分享了 Kubernetes 环境中的安全挑战以及解决方案，为中小企业在 Kubernetes 安全管理方面提供了可复制的经验。

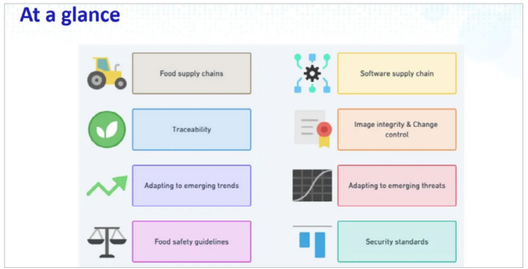


图 1 对比食品和 Kubernetes 环境

2.Kubernetes 环境中的安全挑战

在现代 IT 基础设施中，Kubernetes 已成为容器编排的标准平台，为企业基础设施建设提供了灵活性和高效性。随着 Kubernetes 的广泛应用，其安全性也日益受到关注。Redhat 于 2023 年 9 月发布的《Kubernetes 安全状况报告》表明，安全问题已成为 Kubernetes 部署的主要障碍之一。如图 2 所示，超过三分之二的受访者因安全问题选择延迟 Kubernetes 的部署。

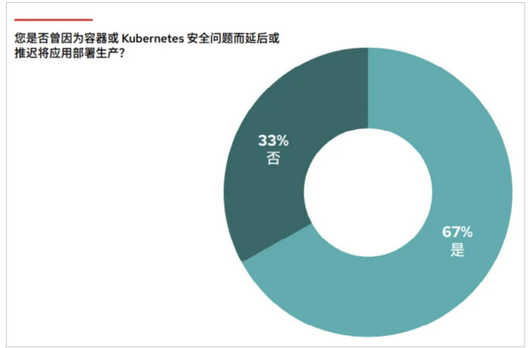


图 2 因安全问题影响 Kubernetes 部署情况统计

导致企业减缓部署 Kubernetes 的安全原因是多方面的，如图 3 所示，Raunaq 和 Caleb 在演讲中从 Kubernetes 威胁矩阵的角度列出了 Kubernetes 环境中的各类威胁，这些威胁主要包括：

- 初始访问

攻击者可以通过利用云凭证或者其他已泄露的凭证来获取对 Kubernetes 集群的初始访问权限。这类攻击的常见方式包括钓鱼攻击、社交工程以及漏洞利用等。

- 执行

一旦获得访问权限，攻击者可以在容器内部执行恶意命令或脚本，以进一步扩展他们的控制。这类攻击可能包括在受感染的容器中运行 shell 命令，或者通过注入恶意代码来执行远程控制。

- 持久化

攻击者往往会尝试通过后门或恶意容器在系统中保持持久化的访问。常见的持久化手段包括安装恶意软件、创建新的用户账号或者修改系统配置以在重启后继续保留控制。

- 特权提升

通过绑定集群管理员角色或其他高权限角色，攻击者可以提升他们在集群中的权限，从而访问和控制更多的资源。常见的手段包括利用漏洞或错误配置来获得更高的权限。

- 防御逃避

为了隐藏他们的活动，攻击者可能会尝试清除容器日志或其他痕迹，以避免被检测和发现。这类防御逃避手段通常包括删除日志文件、覆盖敏感数据以及修改监控配置等。

- 凭证访问

攻击者可能会试图访问 Kubernetes API 服务器以获取更多的敏感信息，如访问令牌、配置文件和密钥等。这类攻击通常通过暴力破解、钓鱼攻击或者利用未加密的通信渠道进行。

- 发现

一旦进入 Kubernetes 环境，攻击者会尝试枚举集群中的资源和服务，以发现潜在的攻击目标。这类发现活动通常包括扫描开放端口、枚举服务和 Pod 以及收集环境配置信息等。

- 横向移动

攻击者可以通过网络在集群内部横向移动，从一个受感染的节点扩散到其他节点。常见的横向移动手段包括利用网络漏洞、滥用凭证以及通过共享资源进行传播。

- 收集

攻击者可能会从受感染的 Pod 中收集敏感数据，包括用户信息、应用数据以及系统配置等。这类数据收集通常通过访问数据库、读取配置文件以及监视网络流量来实现。

- 影响

攻击者的最终目标可能是对系统造成影响，如资源劫持、数据破坏、服务中断等。常见的影响手段包括加密勒索、数据泄露、资源耗尽以及服务拒绝攻击等。

图 3 Kubernetes 威胁矩阵

3. 保护 Kubernetes 环境的关键措施

为了应对上述安全挑战，Kubernetes 环境需要全面的安全保障。然而，正如 Raunaq 和 Caleb 所说，中小企业真正投入业务安全的资源往往是非常有限的，怎么用最少的资源、以最快的速度最有效地发现和修复环境中的安全问题成了当下最棘手的问题。

Raunaq 和 Caleb 分享了一种方案：依托丰富的开源工具减少各关键控制点的安全风险。如图 4 所示，首先需要分析整个软件开发生命周期的关键控制点，然后借助相应的开源工具或技术实施对应的安全策略。这些策略应涵盖容器镜像、配置管理、运行时保护到事件响应的各个方面。接下来笔者将结合他们的观点，就如何保护 Kubernetes 环境进行简要介绍。

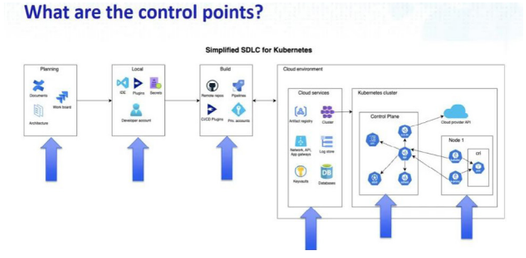


图 4 全软件开发生命周期关键控制点

3.1 镜像安全

- 加固基础镜像

使用 Rapidfort 提供的 Community Images 工具加固基础镜像，选择可信的基础镜像、移除不必要的组件和依赖项。

- 镜像扫描

在镜像构建过程中使用如 Grype 和 Xeol 等工具扫描镜像中的漏洞，减少镜像引入安全漏洞的风险。

- 镜像签名和验证

使用 Notation 对镜像进行签名，可以确保镜像的来源可信，并防止镜像在传输和存储过程中被篡改。

3.2 配置管理

- 扫描错误配置

使用 Kubescape 扫描集群配置（如 YAML 文件和 Helm），并发现潜在的错误配置。

基准配置核查

遵循 CIS 基准，采用 kube-bench 工具对集群进行基准配置核查，确保集群配置的安全性。

3.3 网络安全

- 使用微分段

如图 5 所示，利用 Kubernetes 的 Network Policy 来实现网络分段，限制 Pod、服务、命名空间和集群之间的通信，定义严格的网络策略，减少潜在的攻击面。

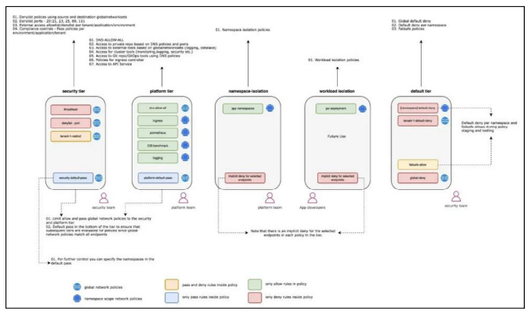


图 5 Kubernetes 环境中的微分段

- 加密通信

使用 TLS 加密服务以及 Pod 间的通信，确保集群内部和外部的所有通信均采用加密协议，防止数据被窃听或篡改。如图 6 所示，在集群中使用 Istio 加密集群网络流量。

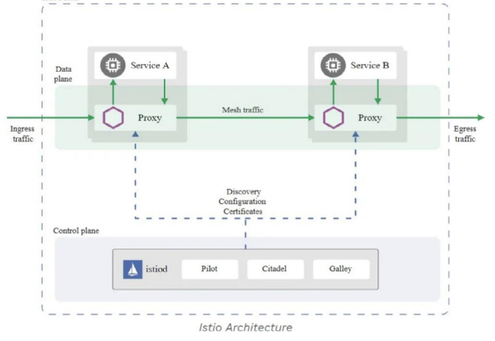


图 6 使用服务网格加密集群流量

3.4 运行时安全

- 运行时监控

使用 Falco 和 Tetragon 等工具进行运行时监控，检测容器内的异常行为和潜在威胁。

- 运行时日志

启用 Kubernetes 的审计日志功能，记录所有的 API 请求和操作，便于事后分析和追踪。

3.5 审计和监控

- 集中日志管理

将所有日志集中管理和存储，便于分析和审计。如图 6 所示，可以使用 Fluentd 来实现集中日志管理。

- 实时监控

部署实时监控工具（如 Prometheus、Grafana 等），及时发现和响应异常行为。

- 报警机制

可以使用 Prometheus Alertmanager 来实现报警通知，当检测到异常行为时及时通知安全人员。

3.6 访问控制

- 最小访问策略

实施 RBAC（基于角色的访问控制），确保最小权限原则。通过定义不同角色的权限，限制用户和服务账户的访问范围，减少权限滥用的风险。

- 权限审计

定期审计和更新权限配置，防止权限滥用。利用 Kubernetes 内置的审计功能，跟踪和记录权限变更和访问活动。

3.7 攻击面管理

- 资产可视化

使用 Cartography 等工具发现集群影子资产以及集群潜在威胁。

- 内部攻击链建模

使用 KubeHound 等工具可视化和分析集群可能存在的攻击路径和漏洞，及时发现内部安全风险。

- 外部攻击面发现

使用 Project Discovery 团队的 PDTM 工具定期进行扫描，及时发现和修复安全漏洞，减少外部攻击。

4. 总结

Kubernetes 作为现代企业容器编排的核心平台，面临着多种复杂的安全挑战。Chipotle 两位应用安全工程师 Raunaq Arora 和 Caleb Schwartz 在本次 RSA 大会上分享了 Kubernetes 环境中常见的安全挑战以及应对策略。笔者认为，随着云原生技术的不断演进，构建安全的 Kubernetes 环境还需要从以下几个方面入手：

4.1 构建基于行为的威胁检测能力

攻击技术日新月异，基于签名的传统威胁检测方法往往无法发现新型攻击，而基于行为的威胁检测系统，仅将已知安全行为作为安全白名单，可以提高对未知威胁的检测和响应能力。

4.2 构建 AI 驱动的异常检测和响应能力

AI 可以通过分析大量数据，识别复杂的攻击模式和异常行为，自动化地进行威胁检测和响应。例如，AI 可以实时分析日志数据，发现潜在的安全威胁并自动触发相应的响应措施，减少对人工干预的依赖，提高响应速度和准确性。

4.3 集成 DevSecOps 能力

随着 Kubernetes 集群的规模和复杂度不断增加，人工管理安全措施变得难以为继，持续漏洞扫描、配置审核和行为监控等安全能力应该被集成到 CI/CD 管道中，以确保在开发和运营的每个阶段都能实时发现并修复安全问题。

总体而言，在 Kubernetes 环境中构建一套全面、智能的安全体系是应对复杂安全挑战的关键。通过多层次的防御措施、零信任模型、自动化和 AI 技术的应用，以及持续的培训和意识提升，才能更有效地保障 Kubernetes 环境安全、稳定运行。

参考文献

[1]https://www.redhat.com/zh/resources/state-kubernetes-security-report-2023.

[2] https://github.com/rapidfort/community-images.

[3] https://github.com/anchore/grype.

[4] https://github.com/xeol-io/xeol.

[5] https://github.com/notaryproject/notation.

[6] https://github.com/kubescape/kubescape.

[7] https://github.com/aquasecurity/kube-bench.

[8] https://juejin.cn/post/7356772031897632818.

[9] https://www.kubernetes.org.cn/9319.html.

[10] https://github.com/falcosecurity/falco.

[11] https://github.com/cilium/tetragon.

[12] https://docs.fluentd.org/v/0.12/articles/kubernetes-fluentd.

[13] https://github.com/kayrus/prometheus-kubernetes/blob/master/alertmanager-deployment.yaml.

[14] https://github.com/lyft/cartography.

[15] https://github.com/DataDog/KubeHound.

[16] https://github.com/projectdiscovery/pdtm.

大模型时代的隐私防护

绿盟科技 创新研究院 陈寅嵩

摘要：当前，大型语言模型（LLM）被广泛运用于各种应用中。然而，这种使用情境下存在一种两难抉择：如何在保护模型所有者的资产和确保用户数据隐私之间取得平衡。在 2024 年 RSA 大会上，来自 Zama 的技术人员 Benoit Chevallier-Mames 与 Jordan Frery 分享了他们如何利用全同态加密（FHE）技术，进一步保护用户与模型供应商的知识产权和隐私。他们展示了这种方法的可行性和实用性，旨在为 LLM 服务提供更加全面的安全支持。

1. 背景信息

随着人工智能技术的飞速发展，大语言模型（LLM）在各个领域的应用日益增多。从简单的写作辅助到复杂的任务，如编辑简历、优化代码，LLM 服务正逐渐成为人们日常生活和工作中不可或缺的一部分。然而，人们使用的 LLM 服务大部分来自于 LLM 服务供应商提供的付费 API。这种方式的访问需要用户将其数据上传到服务端，存在隐私财产泄露的风险，例如著名的某星员工在使用 ChatGPT 时泄露源代码事件。总的来说，LLM 服务存在着一个需要权衡的矛盾，即既需要大模型的能力，也需要对用户的隐私财产进行保护。

针对这个取舍问题，目前有三种主流应对方法：

1.1 使用自研的 LLM

从头开始搭建并训练自己的 LLM 供自己使用固然可以避免泄露用户隐私财产的问题，但是其成本是极其昂贵且难以接受的。这也是为什么目前市场上具备自研 LLM 能力的厂家屈指可数，其他体量较小的用户无力开发自研 LLM。

1.2 使用开源的 LLM

目前开源社区上存在一些开源的 LLM，如 Meta 的 Llama、xAI 的 Grok、国内的 Qwen 等。开源模型的存在有效降低了门槛，那些没有足够计算资源和专业知识的用户也能够轻松地获取、使用和修改这些模型，以解决自己的问题。然而不可否认的是，绝大部分情况下，开源模型的能力普遍不如闭源的商业模型。

Rank	Model	License
1	GPT-4-1106-preview	Proprietary
2	GPT-4-0125-preview	Proprietary
3	Bard (Gemini Pro)	Proprietary
4	GPT-4-0314	Proprietary
5	GPT-4-0613	Proprietary
6	Mistral-Large-2402	Proprietary
7	Claude-1	Proprietary
8	Mistral Medium	Proprietary
9	Qwen1.5-72B-Chat	Qianwen LICENSE
10	Claude-2.0	Proprietary
11	Mistral-Next	Proprietary
12	Gemini Pro (Dev API)	Proprietary
13	Claude-2.1	Proprietary
14	Mixtral-8x7b-Instruct-v0.1	Apache 2.0
15	GPT-3.5-Turbo-0613	Proprietary

图 1 大模型能力排名

1.3 使用商业 LLM

在大部分情况下，用户更倾向于使用能力较强的 LLM 以便更

好地完成任务，因此依然选择使用主流的商业模型，但是这就导致用户不得不将自己的知识财产隐私置于风险之中。

因此，为了实现在使用 LLM 服务的过程中保护用户知识财产隐私的安全这个目标，Zama 提出了在使用商业模型的同时利用全同态加密的方法。

2. 全同态加密

同态加密是一种创新性的加密技术，它提供了一种独特的能力，即在数据加密的状态下进行计算，这种能力对于保护数据隐私的同时允许数据处理至关重要。同态加密的应用范围极为广泛，尤其在云计算和隐私保护的数据处理领域具有巨大的潜力，因为用户可以在不信任服务提供商的情况下，利用云计算资源进行数据处理。同态加密可以分为部分同态加密和层次同态加密，它们分别支持对加密数据进行单一类型的无限次操作（如仅加法或仅乘法）或有限次数的加法和乘法操作。然而，这些形式的同态加密在实际应用中存在一定的限制，因为它们不支持任意复杂的计算。全同态加密（FHE）则是同态加密技术的进阶形式，它允许对加密数据进行任意次数的加法和乘法操作，理论上没有操作次数的限制。

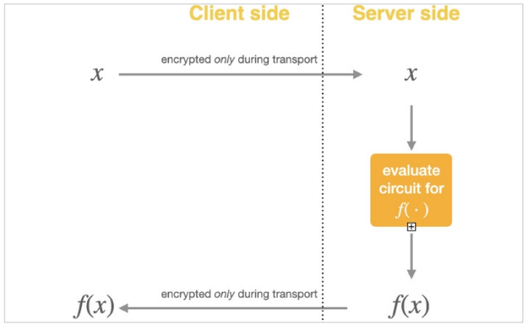


图 2 数据仅仅在传输过程中被加密

当前，在用户使用 LLM 服务时，数据仅仅在传输过程中被加密保护。在服务器侧，用户输入的数据会被解密成明文后再交由大模型处理并生成相应的回复，意味着无论是输入还是输出，数据中用户的所有隐私对于 LLM 供应商来说都是透明的。使用 FHE 后，用户会首先对输入进行加密，模型侧将使用 LLM 对其进行处理得到加密的输出，而用户可以使用自己的密钥对输出进行解密得到输出的明文。全程模型侧都无法接触到输入输出的明文，从而保证用户知识财产和隐私的安全。

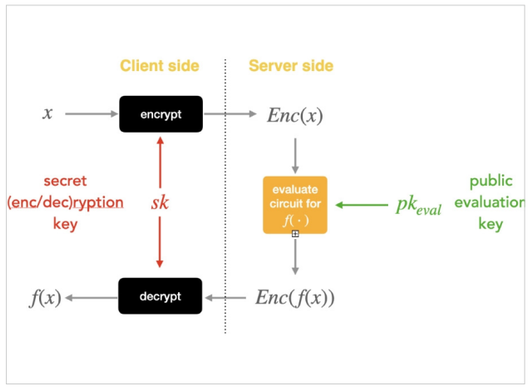


图 3 应用全同态加密后的 LLM 服务

3. 在大模型中利用全同态加密

LLM 所依赖的 Transformer 结构远比普通的加法与乘法要复杂，因此，要分析在 LLM 服务中应用 FHE 的可行性需要对 Transformer 中所有操作进行拆解。总的来说，LLM 主要由三种操作组成：Embedding、Feed-Forward、Attention。

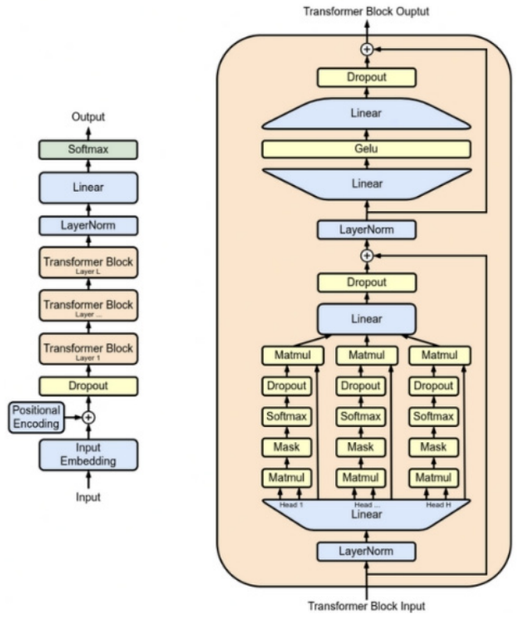


图 4 LLM 的 Transformer 结构

3.1 Embedding 嵌入

在 LLM 中，自然语言输入会通过分词（Tokenization）被分割成一系列更小的单元 Token。为了将这些 Token 转化为模型能够理解和操作的形式，模型会通过查询预定义的词表（Vocabulary）来将这些 Token 映射到唯一的整数 ID。词表是一个包含数万到数十万不等的 Token 及其对应 ID 的列表。每个 Token 在词表中都有唯一的 ID，这个 ID 可以用来检索与每个 Token 相关联的嵌入向量（Embedding Vector）。嵌入向量是模型内部的一种表示形式，它将每个 Token 编码为一个密集的向量，通常是一个固定大小的实数数组，能够体现 Token 在上下文中的意义和用法。

总的来说，LLM 中的 Embedding 操作主要为简单的表格查询操作且不涉及复杂运算，因此对于这部分操作应用 FHE 方案是可行的。通过诸如使用加密后的词表等方式并进行查询，LLM 依然可以快速得到所需的嵌入向量，且并不会花费额外的时间，通常在毫秒级别。

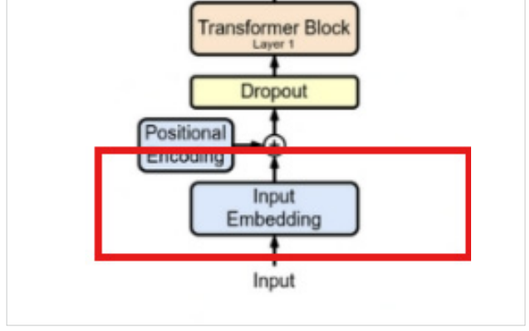


图 5 Embedding

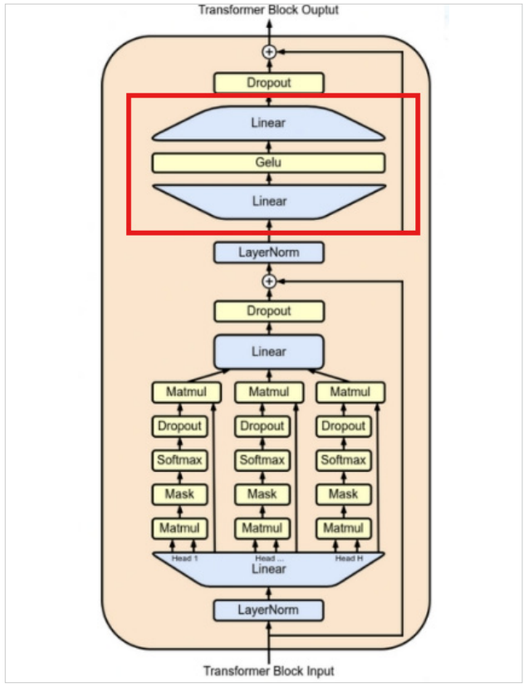


图 6 Feed-Forward

3.2 Feed-Forward 前馈

LLM 结构通常有着数层 Transformer 模块，在每个模块中，注意力（Attention）层后面通常会跟着一个前馈神经网络（Feed-Forward Neural Network），用于接收注意力层的输出，并通过两个全连接层对其进行处理。第一个全连接层将输入维度扩展到一个更大的维度，然后通过激活函数进行非线性变换。第二个全连接层将维度缩减回原始输入维度。这种“扩张—收缩”即是 Feed-Forward 操作，其在不同的表示空间中处理数据进而能够捕获更丰富的特征，起到了增强模型的表达能力和处理局部特征的作用，是 Transformer 模型能够成功处理自然语言理解任务的关键组成部分之一。

从数学上来说，全连接层的操作无非是线性变换，即引入偏置并进行矩阵乘法。全同态加密技术对于密文进行任意加法和乘法的特点可以支持此类型的线性变换。激活函数为非线性变换操作，在全同态加密的背景下，实现非线性函数的加密版本是一个挑战。然而，为了使 FHE 更加实用，研究人员通过选择比较操作、泰勒级数展开、多项式近似等方法，已经实现了一些非线性函数的加密版本，包括激活函数 ReLU、Sigmoid、Tanh。简而言之，FHE 在 Feed-Forward 操作的应用也是可行的，但由于非线性激活函数在加密态下的计算效率可能较低，且为了提升性能 LLM 通常会串联大量的 Transformer 模块导致要计算许多激活函数，其花费时间通常需要超过数秒甚至数十秒。

$$\mathrm{FFN}(x) = F(xW_1 + b_1)W_2 + b_2$$

图 7 Feed-Forward 计算

3.3 Attention 注意力

注意力机制自 2017 年被提出以来，已广泛应用于 LLM 中。它模拟了人类集中注意力的方式，使模型能够捕捉句子中的长距离依赖关系，专注于重要部分，而忽略不重要的内容。在 Transformer 模型中，注意力机制通过“查询 (Q)”“键 (K)”“值 (V)”这三个概念实现。

图 8 Attention

在 Attention 操作中，模型的输入 X 首先经过三次变换，分别得到 Q、K、V。由于这些变换涉及矩阵乘法，属于线性操作，所以可以相对轻松且高效地引入全同态加密（FHE）来获得加密的 Q、K、V。然而，随着如今的模型开始支持越来越长的上下文，Q 与 K 的乘积会变得极其巨大。尽管单次计算不是很久（小于 1 秒），但由于需要大量重复计算，这会对 FHE 的计算性能产生显著影响。Softmax 函数虽然是非线性操作，计算耗时较多，但幸运的是其计算次数相对较少。同样，在与 V 相乘时，巨大的结果也会导致 FHE 性能下降。总的来说，注意力机制在应用 FHE 技术后，其计算时间会大幅增加。仅预测下一个 Token 的操作可能就需要数分钟甚至数小时，而生成完整回复的时间更是以倍数增长。因此，注意力机制是 FHE 在 LLM 领域应用的一个主要性能瓶颈。

$$Q = XW_Q$$
$$K = XW_K$$
$$V = XW_V$$
$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

图 9 Attention 计算

4. 更快的方法——混合模型

为了保护用户的知识财产隐私安全，同时又不损害商业 LLM 服务供应商的利益，Zama 提出了一种混合模型 (Hybrid Model) 的方法。在这种方法中，涉及模型参数较多的操作，如 Feed-

Forward 和 Attention，由服务侧使用全同态加密（FHE）执行。与此同时，其他操作则在用户侧完成。这样一来，大部分的模型参数得到了保护，同时用户的知识产权和隐私也得到了保障。

以 Google 的模型 gemma-7b 为例，其参数分布如下：

Embedding	9.21%
Feed-Forward	71.55%
Attention	15.93%
其他	3.31%

由此可见，权重主要集中在 Feed-Forward 和 Attention 中。因此，在混合模型中，这两部分操作可以在服务侧使用 FHE 执行，而其他操作则在用户侧完成。当然，根据用户的需求，Embedding 操作也可以在服务侧执行。

例如，在一个场景中，用户侧想要向 LLM 进行询问以获得 Attention。用户可以在本地进行分词，并将得到的 ID（整数）使用密钥进行加密后传递给服务侧。然后，服务侧会基于这些密文进行相应的 Embedding 和 Attention 操作，并将密文结果传递回用户侧。用户侧使用自己的密钥解密后，可以选择继续加密并交给服务侧进行 Feed-Forward 操作，或者直接在本地使用自己的 Feed-Forward 神经网络进行下一步操作，而无须再次加密。

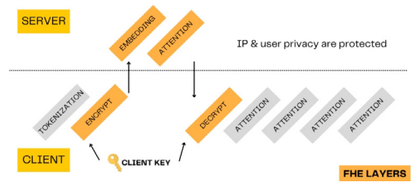


图 10 混合模型

5. 总结

Zama 的技术人员提出了一种创新性的解决方案，利用全同态加密（FHE）技术来保护大型语言模型（LLM）服务中用户和模型供应商的知识财产与隐私，并验证了方案的可行性。这一方案通过引入混合模型的方法，将部分操作放在服务侧执行，同时将其他操作放在用户侧完成，既保护了大部分模型参数，也保障了用户的知识产权和隐私安全。这一突破性的方法为未来 LLM 服务的发展提供了新的方向，有望在保护隐私的同时，推动人工智能技术的广泛应用。

参考文献

[1] https://www.rsaconference.com/events/2024-usa/agenda/session/IP%20Protection%20and%20Privacy%20in%20LLM%20Leveraging%20Fully%20Homomorphic%20Encryption.

[2] https://www.zama.ai/.

[3] https://huggingface.co/blog/encrypted-llm.

工业控制系统网络安全防御体系思考

绿盟科技 售前技术部 王鹏 傅戈 通用领域销售部 马跃强

摘要:随着两化融合的深度推进，工业控制系统在设计、研发、生产、运营、维护等各阶段也不断提升工业化和信息化的融合程度，这也导致工业控制系统在多个环节被攻击的可能性提升。本文在对我国工业控制系统网络安全的现状进行分析后，提出一种集网络安全、全生命周期以及运维与管理体系为一体的多维度的综合网络安全防御体系。

关键词:工业控制系统 安全评估 防御体系 监测预警

1. 引言

当前，我国工业控制系统呈现出数字化、智能化、网络化发展趋势同时也面临着严峻的网络安全威胁。随着两化融合的深度推进，工业控制系统在“设计、研发、生产、运营、维护”等各阶段也不断提升工业化和信息化的融合程度。这也导致工业控制系统在设计、研发、生产、运营、维护等多个环节被攻击的可能性提升 [1]。

2. 工业控制系统工业网络安全现状分析

当前，我国工控网络安全防护工作还处于起步阶段，工业控制系统安全防护意识淡薄，主动开展工作的积极性不高，现有的安全防护手段匮乏，缺乏自我保护能力，大部分工控系统长期处于“亚健康”状态，突出问题表现如下 [2-3]：

（1）缺乏对国外工控产品自主安全控制手段

绝大多数的工业控制系统采用国外知名产品，系统投运时间较长、缺乏维护、升级困难，造成系统普遍存在大量漏洞和后门，对其安全性无法实现自主可控。

（2）工控系统网络防御手段匮乏

工控系统与企业网络、互联网之间缺乏有效的防护策略与措施。工控系统所面临的攻击、病毒、木马等恶意攻击行为已经不再是以往破坏类型，更趋向于窃取、控制的类型，潜伏周期相对较长，如 APT 攻击，需要专业的监测手段与工具，才能及时掌握网络中工控系统的安全状况 [4]。

（3）工控系统安全管理基础薄弱

一是对现场关键工控系统设备与终端保护认识不足。在多数企业，无论是管理人员还是技术人员，不了解工控系统网络安全防护工作的内容，认为工业控制系统不需要保护，使得工业控制系统被攻击路径不断增多，系统安全岌岌可危。

二是现场管理制度体系不健全。未建立专门的工控安全信息管理组织机构，未设置专门管理岗位。现场人员很少接受专门的工控安全培训，缺乏工控系统安全风险识别和应急响应的实战经验。

三是工控网络安全应急管理机制缺乏演练。长期不组织演练，使得应急预案成为一纸空文。

3. 工业控制系统网络安全整体防御体系

基于我国工业控制系统产业长远发展规划以及所面临的网络安全挑战，以工业控制系统关键技术应用和工艺流程特点为落脚点，研究提出一种多维度立体综合网络安全防御体系。

该防御体系贯穿工业控制系统的全生命周期，是结合协议复杂多样、实时性强、节点计算资源有限、设备可靠性要求高、故障恢复时间短、安全机制不影响实时性等特点设计的。采用多层次的纵深协同网络安全防御技术和全方位的监控手段，不断更新和完善技术与管理手段，以实现工控网络可信、可控互联和安全稳定运行。见图 1 所示。

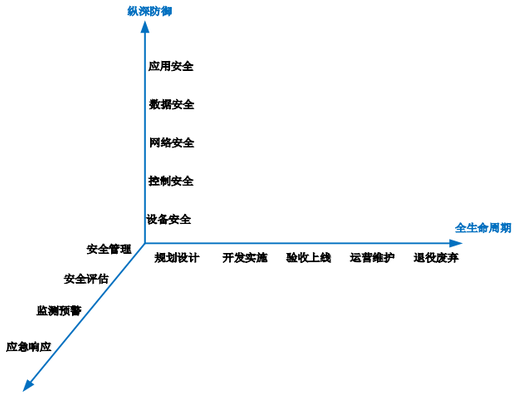


图 1 工业控制系统网络安全防御体系三维架构示意

3.1 纵深协同网络安全防御体系

纵深防御是指基于正确划分工业控制系统控制网络拓扑结构，根据不同区域的特点，从工业控制系统本体、区域、边界等方面采取多层次、系统性、针对性的网络安全防御技术措施。

（1）本体安全防御 [4-5]

本体安全防御 (设备层和控制层) 是指工控设备自身的安全性，针对工业控制系统自动化控制设备的网络安全防护需求，突出自动控制专用网络、协议应用，依托具有深度解析工业属性的网络安全扫描设备，对工控系统进行脆弱性和漏洞早期探测、网络安全风险诊断，针对发现的问题，及时调整安全设计，通过采用补偿加固等终端防护措施，使其漏洞、后门、安全缺陷等隐患得到有效控制；从长远来看，通过持续的技术创新，逐步实现工业控制系统设备 CPU、存储、操作系统内核、基本安全算法与协议等基础软硬件的完整可信、自主可控，未来实现将可信平台植入到工业业务系统。

（2）网络安全防御

在网络安全防御（网络、数据以及应用层）方面，利用工控防火墙、工控专用网络与公共网络间的隔离网关、信息传输加密或敏感数据存储加密、VPN、防病毒、鉴别认证等手段提高工控系统对内外部攻击行为的抵抗能力，并可对系统潜在威胁和风险采取相应的安全措施。例如在生产执行层与管理决策层边界部署身份鉴别、访问控制、入侵检测、行为审计、攻击行为过滤等边界防护措施；在过程控制层与生产执行层边界主要基于工业协议的深度解析，确保过程控制系统与现场控制设备之间、HMI 和现场控制设备之间通信与控制的合法性，通过“黑”“白”名单相结合的防护机制，有效阻止来自办公网的网络病毒、非法入侵、恶意控制等威胁，保障数据、应用安全。

3.2 全生命周期网络安全防御体系

从系统规划、设计、实施、上线、生产、运维到废弃的整个漫长的生命周期中，各个阶段都面临着不同的网络安全问题，必须对工业控制系统网络安全，设计一个适应工控系统特性的全生命周期的网络安全保障体系 , 以持续保障其安全可靠运行。

因此，应当从工业控制系统生命周期的维度，在系统规划、分析、设计、开发、建设、验收、运营和维护、系统废弃的每一个阶段进行网络安全管理，包括：

（1）在系统设计和分析阶段进行安全目标、安全体系、防护蓝图等顶层设计，并将安全防护设计与系统设计相融合；

（2）在系统开发阶段进行代码安全评估，测试阶段同期进行安全测试，包括产品选型、信息系统 IT 产品、工业装备、信息系统安全防护产品和控制系统安全防护产品的安全功能测试和防护能力测试；

（3）在建设完成并验收阶段同时进行风险评估和测评，通过集成测试、协议一致性测试，只有经过网络安全验收测试、风险评估、网络安全保障能力评估后才可上线运行，保障系统安全防护措施的合规性与可靠性；

（4）在运营和维护阶段，应进行周期性风险评估，通过搭建制造工业控制系统攻防环境，进行深入的工控系统漏洞挖掘、攻防演练，及时了解工控系统风险漏洞及攻击手段、路径，持续改进与优化安全技术与管理措施；

（5）在系统废弃阶段做好系统数据的备份和残余信息的销毁等，保障系统保障全生命周期的系统安全。

3.3 全方位网络安全运维与管理体系

严格遵守国家监管政策要求，结合工业控制系统系统的安全防护要素，建立适应工业控制系统网络安全管理与运维体系，从技术、设备、人员、管理、运维等多个维度建立长效安全机制，不断优化改进网络安全防护机制。

（1）安全管理

以风险管理为核心，建立健全工控网络安全管理与运维组织机构，成立管理运维团队，明确安全责任分工，重点从工业控制系统资产安全、软件选择与管理、配置和补丁管理、边界安全防护、数据安全、身份认证、远程访问安全、安全监测和应急处理、供应链管理、落实责任等方面制定建立符合本企业的工业控制系统网络安全管理制度体系，逐渐形成长效的安全机制。

（2）安全评估

针对制造工控系统各层级中的设备、协议、结构、行为和流程等可能存在的威胁，进行专家周期性风险评估，采用科学的风险评估工具和方法，借助全面数据收集、全网攻击路径分析、结构安全性分析、流程审计、网络行为审计等手段，及时发现设备与系统漏洞以及 APT 攻击、DDoS 攻击等网络安全威胁，提出网络安全防护优化与改进方案，实现工业控制系统控制网络的动态安全。

（3）安全监测预警

通过对系统内部的网络流量、文件传输、访问记录等流量、应用和操作行为的综合分析与数据挖掘，分析行为特征，构建行为模型，实现对已知威胁和未知威胁的感知，具备与其他安全防护技术联动和主动防御能力，形成事前预警、事中阻止和事后追溯的管控模式。

（4）运维与应急响应

通过对办公和生产网络内数据流量、网络日志和行为特征的分析，对企业网络异常实现动态分析与监测预警，启动相应应急响应机制，构建主动防御体系，实现持续、动态的安全运维。建立应急响应体系，并加强应急演练，保障应急演练的有效性和可操作性，保障系统能持续运营。

4. 结语

通过分析当前我国工业控制系统网络安全存在的问题以及面临的挑战，提出一种多维度综合网络安全防御体系。该防御体系贯穿工业控制系统的全生命周期，采用多层次的纵深协同网络安全防御技术和全方位的安全监控手段，不断更新和完善技术与管理手段，以实现工业控制网络可信、可控互联和安全稳定运行。

参考文献

[1] 许凤凯 . 智能制造工业控制系统全生命周期信息安全保障 [J]. 自动化博览 ,2016,(01).

[2] 余勇 . 林为民 . 工业控制 SCADA 系统的信息安全防护体系研究 [J]. 信息网络安全 ,2012,(05)：74-77.

[3] 宗健 . 工业 4.0 时代的工控网络安全防护研 [J]. 化工管理 ,2016,(12).

[4] 邱金龙 . 工业控制系统信息安全的未来趋势 [J]. 信息与电脑：理论版 .2016,(4):184-185.

[5] 张敏 , 张五一 , 韩桂芬 . 工业控制系统信息安全防护体系研究 [J]. 工业控制计算机 ,2013,(10).

客户端文件直传对象存储的便捷与安全性探究

绿盟科技 创新研究院 浦明

摘要：客户端文件直传对象存储到底安不安全？本文主要探讨企业租赁云服务商的对象存储服务后，客户端直传文件至对象存储的便捷性及安全问题。

关键词：公有云安全 对象存储 文件上传

1. 为什么选择客户端直传文件至对象存储

典型 C/S 架构下，文件“直传”和“服务端代理上传”方式的上传流程如图 1、图 2 所示：

图 1 服务端代理上传文件至对象存储流程

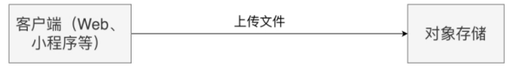


图 2 客户端直传文件至对象存储流程

从以上流程不难看出，服务端代理上传的方式有以下弊端：

（1）网络资源浪费，同一份文件需要上传两次；

（2）服务端带宽消耗大。在云服务器上部署服务端时，考虑到高昂的带宽成本，通常会限制带宽，如果有大量用户上传文件到服务端，可能会迅速耗尽带宽资源，导致请求阻塞，从而影响业务的正常运行；

（3）用户需要自行实现文件上传业务；

（4）客户端到服务端通信的安全性无法保障，由于不同企业的安全能力不同，缺乏统一的鉴权模式，客户端到服务端的通信可能存在风险。

对比服务端代理上传，我们再看看“直传”方式具备哪些优势：

（1）同一份文件只需要上传一次，不会造成额外的网络资源浪费，此外，带宽消耗取决于客户端的带宽大小，因此服务端的压力显著减少。

（2）用户无须自行实现文件上传业务，可将具体操作交由云服务提供商处理（如果上传业务逻辑不复杂，云服务商提供的上传方式基本可以作为参考模板）。例如，阿里云的 OSS 对象存储服务提供多种上传方式 [1]，如图 3 所示：

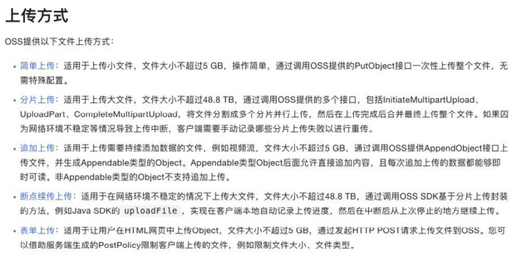


图 3 阿里云 OSS 对象存储文件上传方式

（3）安全性高，云服务商会提供统一的授权模式，如 STS 令牌、第三方签名 URL 等，具体内容可见下文分析。

2. 如何实现客户端直传文件至对象存储

通常情况下，我们使用 Web 端或小程序作为客户端。由于浏览器实施同源策略的跨域限制，所以无法直接将文件上传至对象存储。要实现“直传”，首先需要解决跨域问题。我们可参考云服务提供商提供的跨域设置，以阿里云的 OSS 为例，可以针对特定的存储桶进行相应的配置，如图 4 所置：

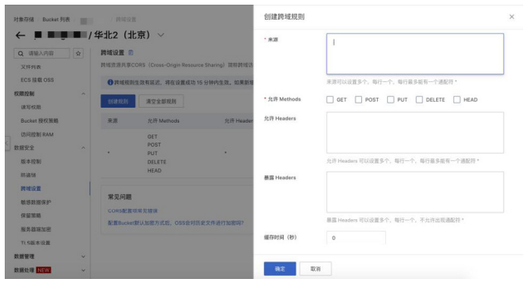


图 4 阿里云 OSS 对象存储跨域配置

解决跨域问题后，将文件上传至对象存储需要获得云服务商的授权。通常情况下，我们可以通过云服务商向用户颁发的 AK/SK 凭证来获取授权。然而，“直传”方式可能会将 AK/SK 暴露在 Web 端页面，存在凭证泄露的风险，因此这种方式并不安全。为降低风险，云服务商提供了多种授权方式，例如阿里云的 OSS 服务提供 STS 令牌、第三方服务签名 URL、表单上传等方式。

3. 客户端直传文件至对象存储的风险分析

3.1 云凭证泄露导致对象存储数据泄露风险

文件“直传”过程中，如果用户在客户端 Web 的 HTML 页面中硬编码云凭证信息，可能会导致云凭证泄露风险。攻击者可以利用泄露的云凭证访问对象存储并窃取数据。云服务商通常提供两种凭证类型：云用户的 AK/SK 和 STS（Security Token Service）临时访问凭证。对于 AK/SK 类凭证，云服务商通常不强制其过期时间，一旦泄露且未定期更换，攻击者可持续窃取数据。对于 STS 临时访问凭证，用户可限制访问权限和有效期，避免长期暴露 AK/SK 的风险。尽管 STS 一定程度上减少了风险，但若用户不遵循最小权限原则或设置过长有效期，仍可能导致数据泄露风险。以阿里云的 STS 为例，STS 凭证只能通过 RAM 角色获取，用户可设置有效期，最长为 12 小时，如图 5 所示：

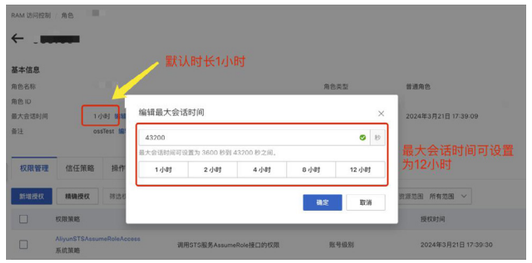


图 5 阿里云 RAM 角色设置 STS 有效时长

即使 STS 的有效时长被设置为最大值，给予了攻击者更多时间，但若 RAM 角色仅被授予有限权限，攻击者仍只能在特定权限范围内窃取数据。若权限设置过大，将扩大攻击范围，导致更多云资源数据被窃取，因此我们可以对 RAM 角色进行授权设置，确保权限范围受控。图 6、图 7 展示了如何为 RAM 角色新增权限以及 STS 临时凭证所支持的云服务。

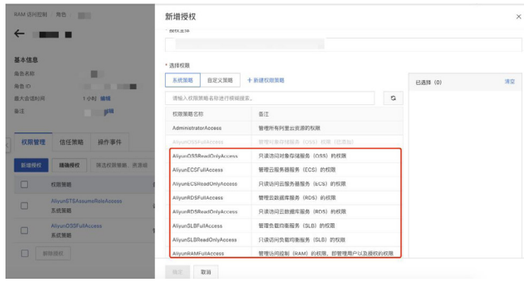


图 6 阿里云 RAM 角色新增授权

云服务	子服务/子模块	RAM代码	控制台	API
云服务器ECS	云服务器ECS	ecs	√	√
块存储	块存储	ecs	√	√
块存储	块存储EBS	ebs	√	√
云服务器ECS	GPU云服务器	ecs	√	√
云服务器ECS	弹性裸金属服务器	ecs	√	√
云服务器ECS	超级计算集群	ecs	√	√
云服务器ECS	专有宿主机	ecs	√	√
云服务器ECS	Alibaba Cloud Linux 2	ecs	√	√
弹性伸缩	-	ess	√	√
容器服务	-	cs	√	√
容器服务Kubernetes版	-	cs	√	√
批量计算	-	batchcompute	√	√
资源编排	-	ros	√	√

支持STS的云服务

图 7 支持 STS 的云服务

3.2 云凭证泄露导致其他云资源被滥用风险

AK/SK 和 STS 云凭证泄露也会带来其他云资源被滥用的风险。攻击者可以使用泄露的云凭证和云服务商提供的 SDK 操作其他云资源，导致资源被滥用。例如，攻击者可以利用云凭证创建 Serverless 函数并进行滥用。云服务商提供的 Serverless 函数通常具有免费试用、低部署成本以及可信的访问域名和非唯一的 IP 等特征，攻击者可将其作为跳板隐藏攻击资产，如图 8 所示：

图 8 Serverless 滥用风险

4. 客户端直传文件至对象存储授权机制分析

客户端直传文件需要对象存储的授权，云服务商提供了多种授权机制以满足用户需求。需要注意的是云凭证不应存储在前端，而应由服务端生成签名或临时访问凭证。我们依然以阿里云的 OSS 举例说明，其提供了三种授权机制：

STS 临时访问凭证机制：服务端生成临时凭证（可设置有效期），返回给客户端。客户端使用 OSS SDK 和临时凭证上传文件，该凭证可重复使用，适用于文件分片传输或断点续传场景。

表单上传机制：服务端生成签名（可设置有效期），主要用于限制客户端上传文件的属性信息，如文件大小、格式、上传路径等，客户端使用签名信息，无须依赖 OSS SDK，以 Form 表单形式上传文件，适用于需要限制上传文件属性的场景。

第三方生成签名 URL：服务端生成签名 URL（可设置有效期），返回给客户端。客户端通过调用签名 URL 上传文件，适用于简单上传场景。

4.1 STS 临时访问凭证机制分析

由于客户端调用 OSS SDK 上传文件时需要携带服务端生成的

STS 临时凭证，所以存在 STS 临时凭证泄漏到前端页面的风险，如图 9 所示：

图 9 客户端调用 OSS SDK 实现文件上传

通过对 STS 临时凭证设置最小权限和最短有效期，可将泄露风险降至最低。最小权限可仅限制为对对象存储的读或写权限，而最短有效期可在后端或控制台中进行设置。 笔者验证了 STS 临时凭证的最短有效期为 15 分钟，如图 10 所示：

图 10 STS 临时凭证最短有效期限

4.2 表单上传机制分析

为了避免前端泄露临时凭证信息，一种可行的方法是使用服务端生成签名的方式。这样，前端就无法通过签名获取具体的令牌信息了。以阿里云 OSS 的服务端生成 PostObject 签名和 Post Policy 为例，服务端生成签名所需的参数输入包括：

PostPolicy 上传策略：用于限制文件属性；

PostPolicy 策略的过期时间：防止被重复使用；

SK 凭证：云凭证信息。

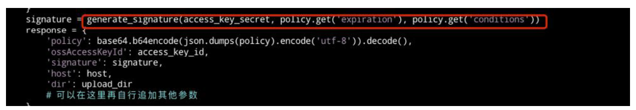


图 11 签名函数参数输入

图 12 签名函数

PostPolicy 文件上传策略是授权机制的核心，通常包含两部分内容：expiration（签名有效期）和 conditions（约束条件）。如图 13 所示：

图 13 PostPolicy 样例

expiration 限制了签名的有效时长，通过 HMACSHA256 哈希算法生成的加密字符串使 SK 凭证泄露风险达到相对可控。

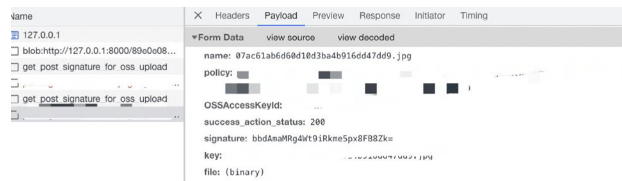


图 14 PostPolicy 签名

conditions 用于限制文件上传属性，防止攻击者利用签名对对象存储进行恶意操作。为了更清晰地探讨此场景，笔者仍以阿里云 OSS 举例说明，假设有以下用户需求：

(1) 文件上传至对象存储 Test Bucket 的“test1/”路径下；

(2) 文件格式为 jpg 或 jpeg 的图片，大小不超过 1MB；

(3) 文件上传请求的 Header 中需要包含 x-oss-meta-prop 的 header key，值为“ssssss”；

(4) 同一文件不能重复上传；

(5) 签名有效期为 1 小时；

用户尝试编写的策略如下所示：

```
policy = {
  # 有效期。
  "expiration": generate_expiration("3600"),
  # 约束条件，参考：https://help.aliyun.com/zh/oss/developer-reference/postobject。
  "conditions": [
    {"bucket": "test" },
    ["eq", "$success_action_status", "200"],
    ["starts-with", "$key", "test1/"],
    ["content-length-range", 1, 1000000],
      ["in", "$content-type", ["image/jpg", "image/png","image/jpeg"]],
    ["eq","$x-oss-forbid-overwrite","true"],
    ["starts-with", "$x-oss-meta-prop", "ssssss"],
  ]
}
```

用户编写的策略应包含上述限制。服务端签名后，将 Policy 信息与 SK 凭证、签名过期时间传递至客户端。客户端通过 Form 表单上传文件。如图 15 所示：

```
<script type="text/javascript">
  const form = document.querySelector("form");
  const fileInput = document.querySelector("#file");
  form.addEventListener("submit", (event) => {
    event.preventDefault();
    let file = fileInput.files[0];
    let filename = fileInput.files[0].name;
    fetch("/get_post_signature_for_oss_upload", { method: "GET" })
      .then((response) => response.json())
      .then((data) => {
        const formData = new FormData();
        formData.append("name", filename);
        formData.append("policy", data.policy);
        formData.append("OSSAccessKeyId", data.ossAccessKeyId);
        formData.append("success_action_status", "200");
        formData.append("signature", data.signature);
        formData.append("key", data.dir + filename);
        formData.append("x-oss-meta-prop","ssssss");
        formData.append("x-oss-forbid-overwrite","true");
        //formData.append("success_action_redirect","http://www.aliyun.com");
        // file必须为最后一个表单域，除file以外的其他表单域无顺序要求。
        formData.append("file", file);
        fetch(data.host, { method: "POST", body: formData })
          .then((res) => {
            if (res.ok) {
              console.log(res);
              alert("文件已上传");
              return;
            }
            if (res.status === 403) {
              alert("缺少OSS操作权限");
            } else {
              alert(`上传请求失败，请求结果：${res.statusText}`);
            }
          })
```

图 15 客户端表单上传

若同一文件上传两次，OSS 会验证策略匹配，并返回 403 无权限。如图 16 所示：

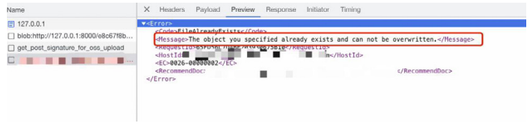


图 16 PostPolicy 策略匹配错误提示

由于有 PostPolicy 限制，该签名调用实际上是一次性的，即使攻击者拿到签名也只能在有限时间内执行一次操作，无法对其他文件做任何操作。即便用户未设置文件覆盖上传，攻击者对 OSS 发起 DoS 攻击，从云服务商责任划分原则上来看，用户自身通常不会承担安全风险，云服务商应具备抵御 DoS 攻击的能力。因此笔者认为该方法可以有效控制临时凭证的泄露风险，并同时保护对象存储系统的安全性。

4.3 第三方生成签名 URL 机制分析

第三方生成签名 URL 机制可通过服务端生成一段具有有效期限的 URL 实现，客户端通过服务端返回的 URL 发起 PUT 请求来上传文件。签名 URL 的输入包括：

上传对象名（Object）；

指定 Headers（用于限制文件属性，如 Content-Type 限制文件类型）；

签名过期时间（expire_time）。

```
def generate_presigned_url():
    # 获取用户传入文件名称
    filename = request.args.get('filename')
    # 填写Object完整路径，例如exampledir/exampleobject.png。Object完整路径中不能包含Bucket名称。
    object_name = 'test1/' + filename
    # 指定Header。
    headers = dict()
    # 指定Content-Type。
    headers['Content-Type'] = 'image/png'
    # 指定存储类型。
    # headers["x-oss-storage-class"] = "Standard"
    # 生成签名URL时，OSS默认会对Object完整路径中的正斜线（/）进行转义，从而导致生成的签名URL无法直接使用。
    # 设置slash_safe为True，OSS不会对Object完整路径中的正斜线（/）进行转义，此时生成的签名URL可以直接使用。
    url = bucket.sign_url('PUT', object_name, expire_time, slash_safe=True, headers=headers)
    return url
```

图 17 签名 URL 输入参数

经过服务端签名后的 URL 格式为：

https://xxxx.oss-cn-xxxx.aliyuncs.com/test1/xxxx.png?OSSAccessKeyId=xxxx&Expires=xxxx&Signature=xxxx

由于 URL 中的 Signature 值经过加密，客户端可以在不透露 AK/SK 访问凭证的情况下，授予第三方在特定有效期内对 OSS 资源的访问权限。此外，客户端无须依赖 OSS SDK 即可实现文件上传。

5. 延伸思考

5.1 对象存储的文件上传流量几乎是免费的，而文件下载流量和存储是付费的

除非使用传输加速功能 [3]，对象存储的文件上传流量通常是免费的。在限制文件大小和上传路径的前提下，如果攻击者获取了用户的云凭证并持续上传文件至对象存储，用户是无须为上传文件的流量付费的，而云服务商需要对攻击流量进行有效防御。

从便捷性和成本效益的角度考虑，用户可尽量避免使用公网流量下载文件，而应该利用云内网通道在同一 Region 内进行文件下载。例如，用户购买的对象存储服务和云服务器位于同一 Region，则可通过内网通道免费将对象存储中的数据下载至云服务器上，进行具体业务处理，最后再上传至对象存储。需要注意的是，同一 Region 的云服务器和对象存储之间内网互通，不同 Region 则不能。

此外，云提供商通常提供对象存储挂载云服务器的功能，如阿里云的 OSS Bucket 可以以目录的方式挂载至 ECS 实例 [4]。

进一步的优化方式可以通过部署 Serverless 函数实现免费文件下载，这种方式的优势在于 Serverless 函数通常具有免费调用额度，并且具备较好的并发能力，可支持多租户上传。

5.2 云租户持续的犯错将导致大量数据泄露事件的发生

过去一年间，云安全领域涌现了一系列重大事件，例如：

2023 年 1 月，Wiz 发现了 Azure Active Directory（AAD）中的一个新攻击向量，影响 Microsoft 的 Bing 服务，该攻击向量基于常见的 AAD 配置错误，使得配置错误的应用程序允许未授权的访问 [6]；

2023 年 2 月 21 日，美国在线新闻网站 TechCrunch 报道称，美国国防部的一个服务器泄露了约 3TB 美国军方内部电子邮件数据。该服务器托管在微软为国防部提供的 Azure 政务云上，理论上该政务云与其他网络是物理隔离的，很可能是错误配置导致邮件服务暴露在了互联网中并且允许匿名访问 [7]；

2023 年 5 月 12 日，Toyota Connected Corporation（以下简称 TC）宣布，丰田汽车公司外包给 TC 公司的部分数据因云环境设置不正确而遭到泄露。此次数据泄露事件影响约 215 万订阅丰田服务 T-Connect、G-Link、G-Link Lite 和 G-BOOK 的用户，泄露的信息包含车辆的位置信息、车辆在上述位置的时间以及车载终

端 ID 和车辆识别号 VIN，甚至包含 2016 年 11 月 14 日至 2023 年 4 月 4 日期间行车记录仪拍到的录像视频[8]；

2023 年 11 月 15 日，Cabernets 披露英国 MPD FM（之前称为 Manpower Direct）使用的亚马逊 S3 (Simple Storage Service) 对象存储由于错误配置泄露了 16,000 多份包含员工护照、签证、身份证、驾驶执照等敏感数据的文件[9]。

从以上事件我们可以看出云租户会持续犯错，或是云服务访问错误配置，或是云凭证被误泄露在了公网上，这些均会导致安全事件的发生。绿盟科技在云上风险发现领域已有多年研究积累，研究发现泄露数据对象的类型较多，典型的如源代码仓库、容器镜像仓库、云数据库、对象存储等。绿盟科技近期发布了《2023 年公有云安全风险分析报告》[5]，报告针对公有云服务配置错误、云服务自身脆弱性配置或漏洞以及云服务的第三方供应链软件三方面进行了分析与总结。此外，绿盟在能力侧也独家发现了全国多个系统源代码暴露情况 —— 超四百万公民个人隐私信息存在泄露风险，这些安全事件均已上报至相关监管机构，从而真正做到了先于攻击者发现云上风险。

6. 总结

本文分析了客户端直传文件至对象存储的便捷性及其面临的安全风险，并探讨了云服务商提供的文件“直传”安全机制的利弊。最后，提出了一些延伸思考，欢迎各位读者批评指正。

参考文献

[1]https://help.aliyun.com/zh/oss/user-guide/upload-objects-to-oss/?spm=a2c4g.11186623.0.0.6b5e6075UQumb8 .

[2]https://help.aliyun.com/zh/ram/developer-reference/api-sts-2015-04-01-assumerole?spm=a2c4g.11186623.0.i155#main-107864.

[3]https://help.aliyun.com/zh/oss/user-guide/enable-transfer-acceleration .

[4] https://help.aliyun.com/zh/oss/user-guide/methods-to-attach-an-oss-directory-to-an-ecs-instance?spm=5176.22414175.sslink.1.303c4e90pjaRdq .

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

[6] https://www.wiz.io/blog/bingbang .

[7] https://techcrunch.com/2023/02/21/sensitive-united-states-military-emails-spill-online/ .

[8] https://company.toyotaconnected.co.jp/news/press/2023/0512/.

[9] https://cybernews.com/security/mpd-fm-passport-data-leak/ .

[10] https://www.zhihu.com/question/461803154/answer/3178042223.

网络安全政策导读（2024年1—6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 1—6 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《关于加强数据资产管理的指导意见》

【内容概述】1 月 11 日财政部发布。针对当前数据资产管理存在的问题，《指导意见》明确要求，以促进全体人民共享数字经济红利、充分释放数据资产价值为目标，以推动数据资产合规高效流通使用为主线，有序推进数据资产化，加强数据资产全过程管理，更好发挥数据资产价值。《指导意见》主要包括总体要求、主要任务、实施保障等三方面十八条内容。

【绿盟观点】早在 2022 年，中共中央、国务院印发《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”），提出了数据基础制度的四大体系：产权、流通交易、收益分配、治理。“数据资产管理”即隶属于其中的“数据流通交易”制度体系。

2023 年以来，国家在数据管理领域陆续启动相关工作，主要包括公共数据价格形成机制研讨、发布《数据资产评估指导意见》、提出“数据基础设施”框架、推进数据基础制度先行区实践、发布《“数据要素 ×”三年行动计划（2024—2026 年）》等。

此次《指导意见》并未将安全单列，而是融入相关管理要求，主要包括四个方面。在标准方面，提出建立数据资产安全标准；在数据资产开发利用方面，提出“原始数据不出域、数据可用不可见”原则，评估运营风险，营造安全可信运营环境；在数据资产过程监测方面，提出分类分级、等级保护制度衔接，数据资产的可追溯要求等;在应急管理方面，提出建立数据资产预警、应急和处置机制。

1.2 16 所高校入选新一期一流网络安全学院建设示范项目

【内容概述】2 月 4 日中央网信办、教育部发布。入选的 16 所高校为：华中科技大学、西安电子科技大学、北京航空航天大学、上海交通大学、山东大学、北京邮电大学、中国科学技术大学、东

南大学、暨南大学、武汉大学、北京理工大学、湖南大学、哈尔滨工业大学、西北工业大学、天津大学、战略支援部队信息工程大学。评选采取“有进、有出”的动态调整机制，每五年一个建设周期，建设周期结束后重新评选。

【绿盟观点】2017 年中央网信办联合教育部发布了《一流网络安全学院建设示范项目管理办法》，随后启动了一流网络安全学院的建设示范工作，并分别于 2017 年、2019 年评选出了两批次共 11 家一流网络安全学院。

此次发布为第三批入选名单。从发布内容和相关情况来看，反映了一流网络安全学院建设示范工作的几个重要变化。一是遴选周期的变化。由之前的每两年一次，改为每五年一次。二是入选名单由累加模式变为更替模式，即实行“有进、有出”的动态调整机制。这一变化在本次入选名单营造反映得比较充分，入选的 16 所高校，有 6 所（暨南大学、北京理工大学、湖南大学、哈尔滨工业大学、西北工业大学、天津大学）是首次入选，其余 11 所皆为第二次入选。三是创新了人才评价机制。将此前的“不唯学历，不唯论文，不唯资历”的评价原则进一步具体化，即不将发表学术论文作为学生毕业、教师晋升的必要条件，而将教师学生参与重要课题、重大工程建设、企业和科研单位技术研发过程中的优秀成果，视同高水平学术论文。

从产、学共建的角度来看，一流网络安全学院建设示范制度的这些改革创新，尤其在网络安全人才评价指标上的重大变革，无疑将为推进网络安全厂商与一流网络安全学院形成更加紧密的联合、聚焦重大关键技术攻关，带来切实的推动力。

1.3《工业领域数据安全能力提升实施方案（2024—2026 年）》

【内容概述】2 月 26 日工信部发布。《实施方案》从总体要求、重点任务、保障措施三方面提出主要内容。在重点任务方面，围绕提升工业企业数据保护、数据安全监管、数据安全产业支撑 3 类能力，明确提出 11 项任务;在保障措施方面，提出了加强组织协调、加大资源保障、强化成效评估、做好宣传引导 4 项工作。

【绿盟观点】政策间的衔接与关联。按照《实施方案》表述，其直接制定依据是《工业和信息化领域数据安全管理办法（试行）》，而从数据安全行业，尤其是厂商的视角来看，《实施方案》与 2023 年年初十六部门联合发布的《关于促进数据安全产业发展的指导意见》相结合，似乎对数据安全产业发展具有直接的指导意义。

《方案》对于促进数据安全产业发展的价值主要体现在三个方面。

一是明确了应用需求的重点方向。《实施方案》对于工业企业数据安全防护能力建设的规划，实际上为数据安全供给行业指明了应用需求的市场。尤其是拟编制的“工业领域数据安全风险防控重点企业名录”、提出的数据安全保护重点场景，以及拟研究制定的“工业领域数据安全保护实践系列指南”，都需要引起足够关注。

二是明确了监管需求的重点方向。监管部门管理手段的建立健全，也是数据安全产业发展的重要市场，在监管体系发展的初期更是如此。《实施方案》所提出的“建设工业和信息化领域数据安全管理平台”，建立“工业领域数据安全工具库”，都需要厂商的技术、产品和方案支撑保障；以“数安护航”专项行动、“数安铸盾”应急演练为代表的数据安全风险防控品牌建设，更是数据安全厂商展示“肌肉”的重要平台。

三是明确了数据安全厂商的能力建设方向。《实施方案》提出的密态计算等关键技术攻关、面向工业云等新应用的数据安全架构设计、“产品 + 服务”供给模式创新等，都是数据安全厂商提升能力的重要参考方向。

1.4《生成式人工智能服务安全基本要求》

【内容概述】3 月 1 日全国网络安全标准化技术委员会发布。该文件规定了生成式人工智能服务安全基本要求包括：语料安全、模型安全、安全措施等，并提出了相应的评估要求。《基本要求》适用于服务提供者开展安全评估、提高安全水平，也可为相关主管部门评判生成式人工智能服务安全水平提供参考。

【绿盟观点】随着 ChatGPT、Sora 等现象级生成式人工智能应用的相继迅速爆火，其应用的安全性日益受到广泛关注。目前，我国针对生成式人工智能的规范性文件以部门规章为主，包括《互联网信息服务算法推荐管理规定》《互联网信息服务深度合成管理规定》及《生成式人工智能服务管理暂行办法》（以下简称《暂行办法》）等。其中，《暂行办法》是我国首部针对生成式人工智能发布的规范性文件，已于 2023 年 8 月 15 日正式实施，它明确了对生成式人工智能服务实行包容审慎和分类分级监管，规定了对生成式人工智能服务提供者的制度要求，初步建立了我国生成式人工智能的基本监管框架。

《基本要求》落实《暂行办法》的相关规定，对生成式人工智能服务提出了四大类安全要求。一是语料安全，包括语料来源、语料内容、语料标注安全等;二是模型安全，包括模型生成内容安全、生成内容准确性、生成内容可靠性等；三是安全措施，包括模型适用人群、场合、用途、服务透明度、内容标识等;四是安全评估，该部分充分衔接并细化了《暂行办法》第十七条“安全评估”的要求。

本次发布的《基本要求》是对 2023 年 10 月公开征求意见稿的修改完善，文件发布形式为“信安标委技术文件”（依据为《全国信息安全标准化技术委员会技术文件制订工作程序（试行）》2021），相较于国家标准，此类技术文件虽不具有标准的约束性，但对于文本成熟度高、适于在全国范围推广的技术文件，可同步推进为国家标准。

对于生成式人工智能服务提供者而言，《基本要求》也具有重要实践参考意义。在监管形势趋严的背景下，相关行业和企业有

必要及时关注《基本要求》内容及动态，为相关安全评估工作提前做好准备。

1.5《促进和规范数据跨境流动规定》

【内容概述】3 月 22 日国家互联网信息办公室发布。《规定》共 14 条，主要内容包括：一是明确重要数据出境安全评估申报标准及免予申报数据出境安全评估、订立个人信息出境标准合同、通过个人信息保护认证的数据出境活动条件；二是设立自由贸易试验区负面清单制度；三是细化数据出境安全评估和个人信息出境标准合同及认证制度，包括对数据出境安全评估的有效期限和延期申请、数据安全保护义务和监督管理责任等。

【绿盟观点】数据和个人信息的跨境流动监管，是数据安全和个人信息保护制度的核心内容之一。《数据安全法》《个人信息保护法》等法律对于数据信息的出境、跨境监管都提出了原则性管理框架。此后，针对数据和个人信息跨境安全管理，国家相关主管部门又相继发布了《数据出境安全评估办法》《个人信息出境标准合同办法》《个人信息出境标准合同备案指南（第一版）》《关于实施个人信息保护认证的公告》等相关规章和规范文件，持续完善和细化监管要求。

《促进和规范数据跨境流动规定（征求意见稿）》曾于 2023 年 9 月面向社会公开征求过意见。从内容变化来看，相关规定更加具体，如对于关基运营者数据出境保护要求和适用条件、数据处理者相关保护和报告义务、监管机构的日常监督重点等，都做了补充和细化。

从《促进和规范数据跨境流动规定》的影响来看，或有三个方面值得关注。其一，进一步厘清和明确数据信息跨境管理的三种基本监管制度（跨境安全评估、个人信息安全认证、跨境标准合同）的适用条件、适用流程等，从操作层面便于数据处理者进行区分和采用。其二，从“减负”的层面，用较大篇幅（第 1—6 条、第 9 条）规定了对于相关监管的例外和豁免、申报便利举措等，进一步限定了数据跨境相关监管制度的范围，并尽量简化监管要求。客观上有利于减轻数据处理者的合规负担。其三，在对相关监管机制的适用条件进行明确和限定的背景下，重要数据、敏感信息的认定等基础问题，无疑又重新回到了数据和个人信息保护管理的“舞台中央”。

1.6《关于深化智慧城市发展 推进城市全域数字化转型的指导意见》

【内容概述】5 月 20 日国家发改委、国家数据局、财政部、自然资源部四部门联合发布。《指导意见》从总体要求、全领域推进城市数字化转型、全方位增强城市数字化转型支撑、全过程优化城市数字化转型生态以及保障措施 5 个方面提出了 13 项指导要求，旨在明确城市数字化转型目标方向，并推进相关工作落地实施。

关于安全方面的要求，《指导意见》明确了四方面内容。一是加强城市数字空间安全管理，健全完善网络安全监测预警和应急处置机制，构建城市网络运行安全管理体系，提升通信网络韧性。二是加快推进城市数据安全体系建设，依法依规加强数据收集、存储、使用、加工、传输、提供、公开等全过程安全监管，落实数据分类分级保护制度，压实数据安全主体责任。三是加强个人隐私保护。四是推进建设有韧性的城市数据可信流通体系，健全数据要素流通领域数据安全实时监测预警、数据安全事件通报和应急处理机制。

【绿盟观点】国家数据局曾于 2024 年 4 月发布《深化智慧城市发展 推进城市全域数字化转型的指导意见》（征求意见稿）。《指导意见》与征求意见稿相比，在基本框架、内容上都有所调整，主要体现在两个方面。一是简化篇幅，突出主体内容。对征求意见稿的第一部分“总体要求”与第五部分“保障措施”作了大幅度精简，只保留“干货”；同时，对于城市数字化发展的个别内容进行了细化，更加具体。二是完善城市数字化安全体系，增加了“加强个人隐私保护”内容，强化了城市全域数字化转型工作对于《个人信息保护法》相关要求的贯彻衔接。

从战略价值看，《指导意见》明确了智慧城市深入发展所必须要解决的两个基本问题。一是明确了数字中国和智慧城市的关系问题。数字中国、智慧城市是我国在不同时期提出的重要发展战略，长久以来对于两者之间的关系，在学术界和产业界都有见仁见智的不同理解。《指导意见》开篇就对此问题给出了简明扼要的回答，即“城市是推进数字中国建设的综合载体”。据此可理解，智慧城市是数字中国建设的区域化呈现，是数字中国目标在城市范畴上的全面落地。二是明确了智慧城市深入发展的路径问题。智慧城市在我国的建设发展已经历了从起步、发展到创新赋能等几个阶段，在数字中国战略提出后，如何持续深入推进智慧城市的高质量发展，成为各界普遍关注的重大课题。《指导意见》提出了“以数据融通、开发利用贯穿城市全域数字化转型建设始终”的发展路径。可见，“数据”将成为智慧城市建设深入发展的关键词，而“城市全域数字化转型”则成为智慧城市建设深入发展新阶段的显著特征。

《指导意见》对于安全的相关要求，集中体现在“（七）提升城市安全韧性水平”一节中，涉及网络安全、数据安全、个人隐私保护、数据可信流通安全四个方面，基本体现了当前监管视角下的城市数字化转型工作中的数字安全关切。这对行业发展尤其是网络和数据安全厂商的市场布局将带来重要启示。一是对于网络和数据安全、个人隐私保护需求，目前市场上有较多产品和方案，厂商关注的重点可能应更多地放在智慧城市的数字化新场景安全需求，以及如何运用新兴技术提升产品和服务的性能及响应程度。二是对于数据可信流通需求，厂商则需要把握和理解数据可信流通体系的核心诉求，即“数据二十条”所提出的“数据来源可确认、使用范围可界定、流通过程可追溯、安全风险可防范”，以此指导部署好数据可信流通体系中的关键安全需求点。

1.7《关于公开征求＜海南自由贸易港国际数据中心发展条例（公开征求意见稿）＞意见的通告》

【内容概述】6 月 26 日中共海南省委网络安全和信息化委员会办公室发布。《发展条例》适用于在海南自由贸易港内开展国际数据中心业务及其相关的监督管理活动，提出海南省人民政府及相关部门应当设立国际数据交易平台和数据出境综合服务平台，强调数据跨境安全，允许国际数据中心业务运营者可以自主选择使用可信、安全的国内外云服务、AI 算力芯片、AI 大模型等开展国际数据中心业务。

【绿盟观点】2018 年 4 月，中央决定支持海南全岛建设自由贸易试验区，此后法规制度建设一直是推进海南省自贸岛建设的关键基础工作之一。2020 年 6 月 1 日，中共中央、国务院印发的《海南自由贸易港建设总体方案》，可视为海南自贸区制度建设的总纲领。其提出了涵盖 11 大类 39 项的制度体系，“数据安全有序流动”成为这 11 大类制度的重要组成部分。

在此后发布的相关法律法规中，数据安全有序流动一直是海南自贸港网络和数据安全制度建设的核心内容之一。如《中华人民共和国海南自由贸易港法》提出“建立安全有序自由便利的数据流动管理制度”、“探索实施区域性国际数据跨境流动制度”；《海南自由贸易港数字经济促进条例》（公开征求意见稿）提出“开展数字身份跨境认证、跨境电子支付、数据制度、数据跨境流动等领域先行先试”等。

本次《发展条例》则从国际数据中心载体建设的层面，具体明确了对保障数据安全、有序流动的要求。一是明确建设两个基础平台，即“国际数据交易平台”、“数据出境综合服务平台”，作为加强数据流动和集中管理的基础设施；二是从上述基础设施功能的角度，明确了“数据分类分级指导、数据出境自评估、数据合规咨询”等数据安全流动职能服务定位。

随着 2025 年前全岛适时启动封关运作日期的临近，预计海南自贸港相关制度法规建设将进一步提速，数据跨境流动安全领域或将出现更多制度创新，海南自贸港也将成为我国数据跨境流动管理制度的重要试验田。

2. 国外篇

2.1 欧盟《网络安全条例》正式生效 [Regulation (EU) 2023/2841]

【内容概述】1 月 8 日欧盟理事会发布。《条例》规定了欧盟实体内部网络安全风险管理、治理和控制框架的具体措施，以及欧盟将设立一个新的机构间网络安全委员会（IICB），并且该《条例》为 CERT-EU 提供了计算机紧急响应小组的扩展授权。下一步，欧盟将按照该条例规定，建立内部网络安全治理流程，逐步实施该条例规定的具体网络安全风险管理措施，并推动 IICB 尽快成立投入运营。

【绿盟观点】《欧盟网络安全条例》与《欧盟网络安全法》[Regulation(EU) 2019/881]、《关于全欧盟网络安全高通用水平措施的指令》（NIS2 指令）、《欧盟信息安全条例》等共同构成欧盟网络安全领域的法律法规框架。这些法律法规除了在效力层级上不同，对于网络安全的规定也各有侧重。

《欧盟网络安全条例》共 6 章 26 条，重点规定了欧盟网络安全高通用性措施、机构间网络安全委员会（IICB）职责、CERT-EU 机构职责、信息共享机制等。

值得一提的是，该《条例》对 CERT-EU 的职能进行了扩展，将其定位为服务欧盟机构、团体的“网络安全服务中心”（Cybersecurity Service for the Union institutions, bodies, offices and agencies），并大幅增加了 CERT-EU 相关的服务项目授权条款，如授权其面向市场有偿开展网络安全广谱监测、漏洞扫描服务等。

2.2 美国《网络安全框架 2.0》（Version 2.0 of Landmark Cybersecurity Framework）

【内容概述】2 月 26 日美国国家标准和技术委员会(NIST）发布。CSF 变化如下：一是适用范围扩大，包括任何部门的所有组织，不仅限于医院和发电厂等关键基础设施。二是“治理”成为重点，包括组织如何制定和执行有关网络安全策略的明智决策，强调网络安全风险是企业风险的主要来源，高级领导者应将其与财务和声誉等其他风险统筹考虑。三是提供工具和指导资源，为特定类型的用户（例如小型企业、企业风险经理和寻求确保供应的组织）提供实施示例和快速入门指南。

【绿盟观点】CSF 是美国 NIST 开发的一套分析和应对网络安全风险的工具集，并通过一系列参考工具（CSF 2.0 Reference Tool）与相应的标准、社区等资源相关联。CSF 旨在为用户提供一种补充性（而非替代性）的网络安全风险管理工具，NIST 还开设了专门网站（https://www.nist.gov/cyberframework），帮助和指导用户更方便地了解和使用 CSF。

CSF 最初的建立依据是第 13636 号总统行政令《增强关键基础设施网络安全性》（Improving Critical Infrastructure Cybersecurity）（2013 年 2 月），该行政令要求 NIST 牵头相关部门和方面，结合现有标准、准则和实践，研究开发“网络安全框架”以减轻关键基础设施的网络安全风险。

自 2014 年发布以来，CSF 共有 3 次大的版本调整，最初版本是 2014 年 2 月发布的《改进关键基础设施网络安全框架》（the Framework for Improving Critical Infrastructure Cybersecurity）1.0 版。其后，2018 年 4 月发布了 CSF 1.1 版，1.1 版与 1.0 版完全兼容，添加了更全面的身份管理和供应链网络安全管理描述。

此次发布为第三个版本，有三个重要调整。一是扩大了覆盖范围。框架 2.0 更名为“网络安全框架”，对象不再局限于以前的

关键基础设施，进一步扩大了适用范围。二是强化网络安全治理。将原来的 5 层功能框架变为 6 层，增加了“治理”（Govern）一层，并作为核心框架要素。三是强调供应链风险管理，提供了更多关于如何评估和管理供应链中的安全风险的内容，突显了供应链安全的重要程度。

2.3《2025 财年总统预算提案》(Budget Proposal For Fiscal Year 2025)

【内容概述】3 月 11 日美国白宫发布。联邦财政年度为每年 10 月 1 日至次年 9 月 30 日。《2025 财年总统预算提案》可自由支配支出共计 7.3 万亿美元，根据不完全统计，2025 财年美国政府预计在网络空间安全投入近 275 亿美元的资金，其中国防网络安全方面约 145 亿美元，非国防网络安全方面约 130 亿美元。2024 和 2023 财年非国防网络安全支出分别为 118 亿美元和 113 亿美元，2025 年非国防网络安全支出预算大幅提升，同比增长约 10.17% 和 15.04%。

【绿盟观点】网络空间安全和发展一向是美国年度预算案的重要关注领域。2025 年度提案在国防领域和非国防领域都对网络空间安全发展的重点方向做出了经费安排。具体如下：

在国防网络安全方面。美国防部将获得 8498 亿美元资金，约 145 亿美元用于网络空间活动支出，包括：网络安全支出为 74 亿美元，网络空间作战支出为 64 亿美元，网络研发支出为 6.3 亿美元。其中网络空间相关预算涉及国防关键基础设施网络安全、供应链风险管理、网络空间关键信息收集分析、工具开发等重点工作。

在非国防网络安全方面。美能源部将获得 4.55 亿美元资金，用于加强人工智能（AI）、网络安全和能源部门的弹性。美卫生与公众服务部（HHS）将获得 13 亿美元用于网络安全实践的支付和激励计划，1.41 亿美元用于保护 HHS 系统和信息安全。美网络安全和基础设施安全局将获得 30 亿美元资金，比 2023 年增加了 1.03 亿美元。美国务院将获得 8 亿美元的信息技术和网络安全资金，较 2023 年增长 17%。美财政部将获得 1.5 亿美元预算用于保护和防御敏感的机构系统和信息，比 2023 年增加 5000 万美元。美国财政服务局（Bureau of The Fiscal Service）将获得 3.96 亿美元资金用于“将所有大型机应用程序现代化并转移到安全云来增强核心政府财务系统的安全态势”，比 2023 年高出 2400 万美元。美司法部将获得 2500 万美元预算用于网络响应和反情报能力，500 万美元用于应对网络威胁。美农业部确保足够的人员配备和关键的信息技术升级，其中包括修复网络安全漏洞等项目，具体资金未说明。美自动化安全办公室将与 NHTSA 研究所合作，解决车辆网络安全风险以及与人工智能相关的风险。

2.4《联邦风险和授权管理计划 路线图 (2024—2025)》(FedRAMP Roadmap 2024—2025)

【内容概述】2024 年 3 月 28 日美国总务管理局发布。《路线图》概述了 2024 至 2025 年的四个战略目标：第一，以客户体验（CX）为导向。第二，将 FedRAMP 定位为网络安全和风险管理领域的领导者。第三，扩大值得信赖的 FedRAMP 市场的规模和范围。第四，通过自动化和技术前沿运营提高计划效率。

【绿盟观点】为了实现信息技术现代化和云安全的战略目标，美国政府于 2010 年在《改革联邦政府 IT 管理的 25 条实施计划》中提出“云优先”（CloudFirst）策略，（后更新为“云智能”CloudSmart）战略，次年 12 月，美国启动“联邦风险与授权管理计划”（FederalRiskandAuthorizationManagementProgram，简称 FedRAMP），是一项政府范围具有“强制性”的计划，用于标准化的云产品和服务的安全性评估、授权和监控，成为根据《联邦信息安全现代化法案》（FISMA）制定的首个政府层面的安全授权计划。后续又出台了《FedRAMP 授权法案》《关于联邦风险授权管理计划现代化的备忘录》等一系列配套法规和政策。

FedRAMP 的管理机制包括两部分。一是云计算服务监管机制。包括行政管理和预算局（OMB）负责协调各机构“云智能”战略的执行和联邦政府云计算采购机制的运行；国土安全部负责威胁通知协调、事件响应报告等。二是云计算服务审查机制。包括联合授权委员会（JAB）和 FedRAMP 项目管理办公室（PMO），负责制定安全基线要求、对云计算服务进行安全评估、授权、持续监督等，JAB 由国防部（DoD）、国土安全部（DHS）和总务管理局（GSA）构成，是主要的决策机构。美国国家标准与技术研究院（NIST）在其中负责制定联邦政府安全采用云计算服务的提供标准和规范指南等。截至目前，美国 FedRAMP 共授权 400 多家云计算服务提供商（CSP）以及 40 多家第三方评估机构（3PAO）。

本次《路线图》在实施策略方面具有三个较为明显的特点。一是推进数据和服务共享，整合和优化数据中心基础设施，通过数字化的集中服务来降支提效。如发布“数字授权包”、持续诊断与缓解（CDM）仪表板集成等。二是加强与网络安全和基础设施安全局（CISA）合作，推动私营企业参与。如将 CISA 安全云业务应用程序（SCuBA）指导纳入安全配置文件、与 CISA 合作进行红队和专门审查等。三是支持多云环境，增强 FedRAMP 系统的安全保障能力。如定义 FedRAMP 的核心安全期望、发布 FIPS140 的更新指南等。

我国对于云计算安全也建立了相应的评估授权机制。2019 年 7 月，网信办等四部委联合发布《云计算服务安全评估办法》，并于 2020 年启动了评估工作，旨在“提高党政机关、关键信息基础设施运营者采购使用云计算服务的安全可控水平”。该评估由国家市场监管总局下属的“中国网络安全审查认证和市场监管大数据中心”（简称网数中心）牵头组织开展，具体评估工作由 8 家经认证的专业机构承担。

2.5 美国国家标准和技术委员会启动“人工智能风险和影响评估”计划 (Assessing Risks and Impacts of AI)

【内容概述】5 月 28 日美国国家标准和技术委员会发布。“人工

智能风险和影响评估”计划（ARIA）旨在开发一套新的方法和指标，以量化、评估人工智能风险和影响，包括 3 个评估级别：一是模型测试，以确认其能力；二是红队测试，以进行压力测试；三是现场测试，以调查人们在日常使用中是如何接触人工智能的。初步评估（ARIA 0.1）将作为试点工作进行，重点关注与大型语言模型（LLM）相关的风险和影响。

【绿盟观点】美国在人工智能治理监管方面先后发布过多项政策法规。包括：《人工智能权利法案蓝图》（白宫，2022 年 10 月）、《人工智能风险管理框架 V1.0》（美国国家标准技术研究所（NIST），2023 年 1 月）、《关于安全、可靠和值得信赖的人工智能开发和使用的行政命令》（白宫，2023 年 10 月）等，这些政策法规主要旨在提出人工智能监管的原则框架。

美国 NIST 此次推出的“人工智能风险和影响评估”计划（ARIA），主要是对《关于安全、可靠和可信的人工智能的总统行政命令》相关要求（“发起一项举措，制定指导和建议评估和审核人工智能能力的基准”）的响应，并对此前发布的《人工智能风险管理框架 V1.0》中的风险衡量功能进行了扩展，开创了对人工智能安全监管的测评机制。

从内容来看。ARIA 开发一套新的方法和指标以量化、评估人工智能风险和影响，其主要内容包括：一是确定并采用适当的方法和指标，ARIA 建立了模型测试、红队测试、现场测试 3 个不同评估级别，并确定了不同的方法和指标（表 1）；二是评估系统的可信特征，包括日志收集、签署参与协议等；三是完善特定风险识别跟踪机制，包括开发代理任务、推出“测试包”（TPs）等；四是定期评估和反馈测量功能的效果，包括发布 ARIA 0.1 试点评估报告等。

从管理机制来看。ARIA 的评估任务由 NIST 的信息技术实验室（ITL）负责开展，同时，美国人工智能安全研究所（AISI）为大规模增强和生成 ARIA 评估提供协助。

从评估结果的应用来看。ARIA 评估结果公开发布，并且基于此预计将形成包括评估政策、工具、方法和指标在内的系列成果。这些成果一方面将用于评估系统安全性，另一方面还将成为设计、开发、部署或使用人工智能技术的参考。

表 1 不同级别人工智能系统的评估方法和指标

评估级别	评估方法	评估指标
模型测试	对测试包定义的任务和系统行为范围进行评估测试输出	每个查询-响应交互的百分比和/或用户会话的百分比
红队测试	收集每个应用程序尽可能多的违规结果、模型护栏和安全机制进行压力测试	对抗性交互的分析即为分数汇总
现场测试	人类参与者在测试或控制条件下跨多个会话与真实设置中的 AI 应用程序进行交互	指标为评估员关于 LLM 操作的输入、现场测试人员的反应，应用感知与应用实际输出的一致程度即为分数

目前，我国对人工智能安全监管相关的法规政策主要有《生成式人工智能服务管理暂行办法》、《网络安全技术 生成式人工智能服务安全基本要求》等，这些政策仅对于“具有舆论属性或社会动员能力的互联网信息服务”建立了安全评估的机制，而对于生成式人工智能安全的监管，只是在相关的服务监督检查中提及安全评估，但如何开展此类评估，尚未建立具体的制度要求。美国“人工智能风险和影响评估”计划的推行，或对我国后续建立健全生成式人工智能安全评估制度，具有一定的参考作用。

《工业领域数据安全能力提升实施方案（2024—2026年）》指导下的企业数据安全防护探索

绿盟科技 总体技术部 杨博 曹雅楠

摘要：工业数据在工业数字化转型过程中扮演着核心角色。它不仅是工业企业实现数字化建设的关键资源，也因其流通和投资价值而变得日益重要。本研究探索了《工业领域数据安全能力提升实施方案（2024—2026 年）》指导下的企业数据安全防护体系建设，希望在工业互联网基础设施建设和工业信息安全防护实施的关键阶段，能为工业企业、地方监管机构、工业互联网平台运营商提供数据安全建设借鉴，共同确保不同级别工业数据的有效管理和安全保护。

关键词：工业领域 数据安全 分级保护 流转监测

1. 工业领域数据安全法律政策体系日趋完善

在《网络安全法》颁布施行四年多以后，数据安全也成为国家网络安全的重要议题。随着《数据安全法》颁布施行，2021 年成为我国的数据安全立法元年，数据安全保护工作上升到国家法律层面，这是对工业领域数据安全的一次全面规范和提升。

《工业数据分级分类指南（试行）》（工信厅信发〔2020〕6 号）、《关于组织开展工业领域数据安全管理试点工作的通知》（工信厅网安函〔2021〕295 号）等政策文件经过几年的试点工作，已经逐步形成了工业领域数据安全可复制可推广的管理模式和建设经验，促进提升了工业领域数据安全保护水平。

基于工业领域数据安全已经取得的建设成果，2022 年 12 月，为了推进《数据安全法》在工业和信息化领域具体实施，促进工业数据资源的风险防控，工业和信息化部出台了《工业和信息化领域数据安全管理办法（试行）》（工信部网安〔2022〕166 号），为工业领域提供了明确的数据安全管理指导和操作框架，将更好地指导构建工业领域数据安全管理体系，督促工业企业落实数据安全主体责任，加强数据分类分级管理、安全防护和安全监测等工作，有助于提升整个工业领域的数据安全保护水平，确保数据安全和促进数据的合理利用。

为了将数据安全法律政策在工业领域再细化、再落实，提供具有指导性和可操作性的具体措施，加强试点成果转化应用，2024 年 2 月 23 日，工业和信息化部印发《工业领域数据安全能力提升实施方案（2024—2026 年）》（工信部网安〔2024〕34 号，以下简称《实施方案》）。

2. 指导企业构建坚实的工业数据安全防护体系

《实施方案》的第一项重点工作是提升工业企业数据保护能力，

从数据安全意识、重要数据保护、数据安全管理和重点场景数据安全保护四个方面落实工业企业数据安全主体责任。

图 1：工业企业数据保护能力

基于《实施方案》，工业企业可以从以下五个方面提升工业数据安全保护能力。

（一）以分类分级奠定工业数据防护体系基础

企业开展工业数据分类分级、识别报备重要数据，是进行数据安全建设的第一步。首先，企业结合自身的生产制造模式，依据《工业数据分类分级指南（试行）》等相关政策文件，制定适合业务特点的工业数据分类分级模板；其次，依据数据遭篡改、破坏、泄露或非法利用后可能带来的潜在影响，按照《工业领域重要数据和核心数据识别规则（草案）》明确保护对象和级别；最后，通过“自动化工具 + 专家服务”的模式，对数据所属类别与级别进行标记，对已经标记的数据做人工校验审核，形成分类分级数据目录，识别重要数据和核心数据。同时，根据数据分级管控原则，将分类分级结果与数据标签相结合，在工业数据流转及使用环节依托数据标签所定义的数据安全级别，为数据安全防护手段的建立奠定基础。

（二）以工业数据分级保护实现业务与安全融合发展

根据工业生产各环节不同的安全需求，与数据的分类分级结果进行有效关联，针对工业数据在采集、传输、存储、处理、删除、销毁各生命周期阶段不同的安全风险进行差异性防护，采取数据加密、访问控制、数据审计、数据防泄露、数据脱敏等多种保护措施，有效应对数据安全风险。对于未授权数据处理，实现数据拿不到、看不懂、改不了、赖不掉。

（三）以风险评估促进数据防护体系动态应对安全风险

为了实现工业数据安全保护措施与安全需求相匹配，精细化落实数据分级保护的要求，需要定期实施工业数据安全风险评估工作。依据数据安全技术与管理标准，定期评估工业数据资产保密性、完整性和可用性等安全属性和面临的威胁，以及威胁利用脆弱性导致工业数据安全事件的可能性。同时，结合安全事件所涉及的工业数据资产价值来判断安全事件一旦发生可能对工业生产造成的影响。通过识别工业企业内部存在的数据安全风险与外部存在的潜在威胁，减少由于内部管理或外部非受控风险而造成的敏感数据泄露。

（四）以流转监测构建数据流转安全屏障

工业数据在企业生产经营流转过程中的安全问题轻则造成直接经济损失，重则导致生产安全事故。因此，保障工业数据在流转过程中的安全性，是工业企业必须重视的一项工作。以工业数据流转监测平台为中心，采用数据流转探针旁路镜像模式部署，接入工业互联网平台，工业企业互联网、管理信息网、工业控制网的安全日志和流量日志，进行实时数据分析，对工业数据资产的交换、共享、操作等数据流转情况进行安全监测，将工业数据流转异常行为及漏洞等安全风险及时通报给企业相关人员，并具备向上级平台报送的能力，协助构建国家工业数据流转安全综合监测体系。工业数据流转监测技术架构如图 2 所示。

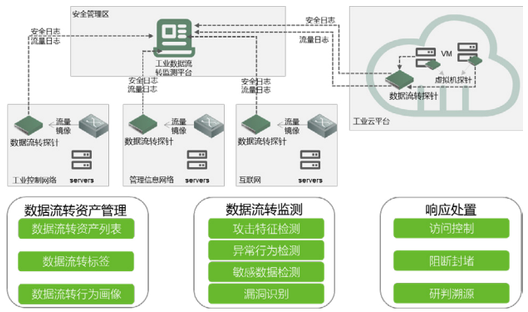


图 2：工业数据流转监测技术架构示意图

（五）支撑工业领域数据安全三级监测体系建设

《实施方案》明确要求强化“部—省—企业”工业数据安全监测技术三级联动，不断提升保障工业领域数据安全的能力，旨在通过分层级的监测策略，打造工业领域数据安全态势全局化汇聚能力，实现对工业数据安全的多维度、全周期保障。

通过安全工具流量采集、云端监测、主动上报等模式，实现对数据资产识别、数据流转监测、数据违规泄露发现和数据安全态势的有效掌握。工业数据安全监测平台技术架构如图 3 所示。

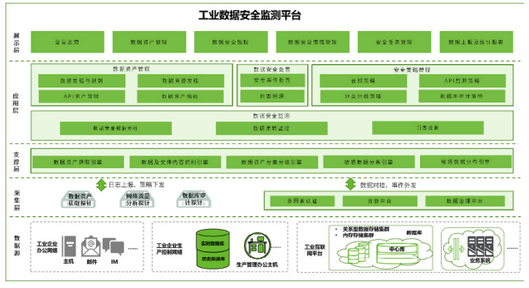


图 3：工业数据安全监测平台技术架构

工业企业建设工业数据安全监测平台，通过大数据处理与分析实现多元化异构数据的接入、管理、存储、运算和智能整合，提供丰富的安全数据分析算法模型，全面提升工业数据安全监测能力。通过构建工业数据安全监测平台，以工业企业为核心、行业主管部门遵循体系化和系统化思维，加强安全资源储备，帮助工业企业及时发现数据安全短板，提升工业数据安全保障能力。

工业数据风险通报基于汇集到的各类数据安全事件、威胁情报信息、安全监测信息、人工录入信息进行综合分析，建立风险通报模版，完成对工业企业数据安全情况的综合分析，生成通报文件。对上可为国家级平台提供数据风险情况，对下可服务于工业企业有效应对风险、加强数据防护支撑。

3. 工业企业数据安全保护体系实施部署

工业数据分级保护框架部署如图 4 所示。

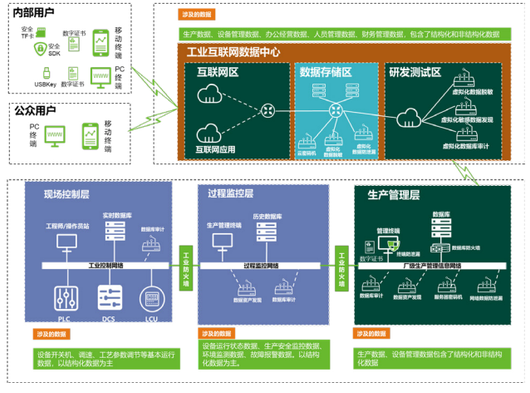


图 4：工业数据分级防护部署示意图

（一）针对工业生产控制网络中的数据安全分级保护部署如下:

在企业工业生产控制网络的现场控制层部署数据库审计设备；在过程监控层部署数据资产发现设备和数据库审计设备；在生产管理层部署数据库防火墙、数据库审计、数据资产发现、服务器密码机、网络防泄露、数字证书、终端防泄露设备，通过上述设备，对生产控制网络中的数据依据分类分级结果实现相应安全等级的防护。

（二）针对工业互联网平台的数据安全分级保护部署如下：

在企业工厂侧与工业互联网平台侧的网络通信中采用专线、VPN 通道以及重要数据签名、加密等方式实现安全传输。

在工业互联网平台的研发测试区部署虚拟化的数据脱敏、敏感数据发现和数据库审计设备。

在工业互联网平台的数据存储区部署云密码机、虚拟化数据脱敏和数据库审计设备。

（三）通过互联网访问工业互联网平台时，针对不同的用户类别，实现不同的数据访问保护级别。

（1）对于内部用户互联网实现远程访问

使用移动终端，需要为移动终端配置安全 TF 卡或者安全 SDK，结合数字证书，实现互联网数据访问的分级保护；

使用 PC 终端通过互联网实现远程访问，需要配备 USBKey 结合数字证书，实现互联网数据访问的分级保护。

（2）对于公众用户

将经过脱敏后的一般数据向公众用户开放，实现公众用户数据访问的分级保护。

对于工业数据安全分级保护，依托工业数据安全监测平台提升事前预防、事中感知、事后审计及追查的综合防控能力，对工业数据产生、传输、存储、共享、处理、销毁全生命周期进行安全管控。

4. 结语

随着信息技术的飞速发展，国家对于数据安全的重视程度日益提高，相应的法律法规及政策文件也在不断完善。在这一背景下，工业数据安全保障能力的建设显得尤为迫切。工业企业正加速向数字化转型，通过引入智能制造、物联网、大数据分析等技术，提高生产效率和产品质量，导致工业数据量激增。随着技术的进步，安全威胁的形式也在不断变化。从传统的病毒、木马到现在的高级持续性威胁 (APT)、零日漏洞攻击等，对工业企业的数据安全构成了严峻的挑战。这些因素都要求工业企业高度重视数据安全保障工作，从技术、管理、人员等多个层面共同努力，同时也需要国家政策的支持和社会的广泛参与，为企业的可持续发展提供坚实的数据安全保障。

参考文献

[1] 刘晓曼 , 冯开瑞 , 郭茜 .《国内外工业领域数据安全发展浅析与趋势展望 》[J]. 保密科学技术 , 2022(03): 21-25.

[2]《工业数据分级分类指南（试行）》（工信厅信发〔2020〕6 号）

[3]《关于组织开展工业领域数据安全管理试点工作的通知》（工信厅网安函〔2021〕295 号）

[4]《工业和信息化领域数据安全管理办法 (试行)》（工信部网安〔2022〕166 号）

[5]《工业领域数据安全能力提升实施方案（2024—2026 年）》（工信部网安〔2024〕34 号）

关于GB/T 43697-2024《数据安全技术 数据分类分级规则》

绿盟科技 业务服务产业部 曾令平

摘要：2024 年 3 月 21 日，全国网络安全标准化技术委员会发布《GB/T 43697-2024 数据安全技术 数据分类分级规则》（以下简称《数据分类分级规则》）国家标准的报批稿，规定了数据分类分级的原则、框架、方法和流程，并给出了重要数据识别指南。2024 年 4 月 17 日，国家标准全文公开系统公布了国标最终版，并于 2024 年 10 月 1 日实施。

1. 概述

数据分类分级最初是网络运营者对数据资产进行一致性、标准化管理的方法，随后成为网络数据安全风险管理的技术方案。《网络安全法》第 21 条的规定：网络运营者应当采取数据分类、重要数据备份和加密等措施。《数据安全法》第 21 条也提出：国家对数据实行分类分级保护。由此可见，数据分类分级的主体由“网络运营者”转变为“国家”，而发布的《数据分类分级规则》正是基于以上上位法的法规要求确立了基本的数据分类分级规则、框架及方法等，是第一部以“数据安全技术”命名的国家标准，是指导各行业、各领域数据分类分级工作的基础性国标，同时也彰显了数据分类分级工作是数据安全的第一步。因此，本文通过阐述对以下三方面、九个点的理解，希望能为组织数据分类分级工作提供一些参考。

2. 提供了基础框架

《数据分类分级规则》总共分为 7 个大章和 10 个附录，覆盖范围、术语与定义、基本原则、数据分类规则、数据分级规则、数据分类分级流程、数据分类分级参考等内容，可为各类组织（如行业领域主管部门、各个数据处理者等）提供详实的参考。如图 1 所示。

图 1 数据分类分级基础框架

3. 解决了什么问题

3.1 对数据分类分级工作职责进一步明确

在《数据分类分级规则》第 5.1 节中，在描述数据分类的业务属性时，提到了“责任部门”这一关键属性，即按照数据管理部门或职责分工进行细化分类。这一条，指定了数据管理部门作为数据分类分级工作的主责部门，也给数据分类分级实践工作提供了一个很明确的职责划分指引。众所周知，在大多数组织实际工作中，数

据分类分级工作的职责划分一般有以下三种：

(1) 由安全(或科技) 部门作为主责部门,牵头数据分类分级工作;

(2) 由数据管理部门 (或数据中心) 作为主责部门，牵头数据分类分级工作 ;

(3) 由安全部门和数据管理部门共同承担，安全部门承担管理职责，数据管理部门承担主导职责。

由于数据与业务强耦合，而数据管理部门又承担了组织数据治理工作职责，而将数据分类分级工作的主要职责放在数据管理部门，似乎是理所应当的。职责有效确定是落实数据分类分级的关键步骤，所以本条内容的确立，为组织落实数据分类分级工作提供了强有力的落地指引。组织可在已发布或制定中的《数据安全管理办法》或数据分类分级相关管理制度中调整或增加数据分类分级相关职责。

3.2 数据主体分类属性类比数据确权授权机制

2022 年 12 月 19 日，中共中央、国务院印发《关于构建数据基础制度更好发挥数据要素作用的意见》(以下简称“数据二十条”)，“数据二十条”指出“建立公共数据、企业数据、个人数据的分类分级确权授权制度”，并就公共数据、企业数据、个人信息的确权授权机制进行了专门说明。

在《数据分类分级规则》第 5.1 节中，提及了另一个关键的业务属性——“数据主体”,虽然本标准是从数据分类的角度出发，但不难发现数据主体的分类思想与“数据二十条”中的数据确权授权机制有着异曲同工之妙，即在数据分类过程中确立了各参与方的数据权属。这也为数据要素化、数据资产化、数据价值化提供了理论基础。

3.3 针对数据集提出了可落地实践的分级思路

在《数据分类分级规则》第 6.2 节中，对“数据集”这一名词做了解释：是由多个数据记录组成的集合，如数据库表、数据库一行或多行记录集合、数据文件等。现阶段数据分类分级实践中，大多是针对结构化数据进行数据分类分级：有的组织细化到对数据项 (通常表现为数据库表某一列字段) 进行分类分级并进行后续分级管控 ; 有的组织针对数据表进行分类分级，即同样先对数据库表中的数据项进行扫描并进行分类分级，不过后续则依据就高不就低原则对数据库表、一行或多行集合等进行分级管控。

数据集，其实更多地为非结构化数据提供了可落地实践的数据分类分级之路，如上可对应结构化数据中的数据库表的分级管控思路。那么，针对常见的非结构化数据包括短信、邮件、图片、音频、视频、办公文档、演示文稿等的分类分级，如一张图片可将其视为数据集，对数据集中的每一个数据项进行识别后并进行分类分级。其中，非结构化数据的分类，其实在现有的数据分类框架下，已然包括了非结构化数据，如没有覆盖的，可参考数据分类方法中业务属性的方法进行新增分类 ; 而数据分级则可按照就高不就低原则，将数据集包含数据项的最高级别作为数据集默认级别。[可参见 6.6 节中第 d) 条内容]

3.4 级别确定规则对于重要数据判定留有余地

在《数据分类分级规则》第 6.4.1 节中，影响对象，是指数据面临安全风险时可能影响的对象。影响对象通常包括国家安全、经济运行、社会秩序、公共利益、组织权益、个人权益。在重要数据的定义中，通常情况下只包括了国家安全、经济运行、社会秩序、公共利益这几个影响对象。而一般数据的影响对象则包括了组织权益、个人权益。如表 1 所示。

影响对象	影响程度		
	特别严重危害	严重危害	一般危害
国家安全	核心数据	核心数据	重要数据
经济运行	核心数据	重要数据	一般数据
社会秩序	核心数据	重要数据	一般数据
公共利益	核心数据	重要数据	一般数据
组织权益、个人权益	一般数据	一般数据	一般数据
注：如果影响大规模的个人或组织权益，影响对象可能不只包括个人权益或组织权益，也可能对国家安全、经济运行、社会秩序或公共利益造成影响。			

表 1 数据级别确定规则表

需要注意的是表中备注说明，也为影响到大规模的个人或组织权益的一般数据确立为重要数据提供了一个参考依据。特别是在工业、汽车等领域，考虑到行业特殊性，则将组织权益或个人权益作为影响对象之一。

(1)《工业和信息化领域数据安全管理办法》: 危害程度符合下列条件之一的数据为重要数据——造成重大数据安全事件或生产安全事故，对公共利益或者个人、组织合法权益造成严重影响，社会负面影响大 ; 引发的级联效应明显，影响范围涉及多个行业、区域或者行业内多个企业，或者影响持续时间长，对行业发展、技术进步和产业生态等造成严重影响。

(2)《汽车数据安全管理若干规定 (试行)》: 重要数据是指一旦遭到篡改、破坏、泄露或者非法获取、非法利用，可能危害国家安全、公共利益或者个人、组织合法权益的数据。

虽然,仅影响组织自身或公民个体的数据一般不作为重要数据，但考虑行业的特殊性以及数据规模等因素，重要数据的影响对象仍可能包括个人、组织合法权益，或在实际中，这类一般数据不被判定为重要数据亦不用上报，但仍需按照重要数据的保护要求进行保护，这在电信行业亦有相关实践。

3.5 细化了个人信息的分类以及典型数据示例

随着业务的发展和具体实践 (如司法判例、行业标准等) 不断加深，大家对个人信息的理解也逐渐清晰。《数据分类分级规则》附录 B 中，在 GB/T 35273—2020 的基础上，细化了个人信息分类，增加了个人信息典型数据示例。具体说明如下 :

(1) 典型数据示例进行了大量扩展，如个人基本资料、个人身份信息、个人生物识别信息、网络身份标识信息等。特别地，部分数据示例与日常生活或实践结合比较紧密，如个人身份信息类中的“证件照片或影印件”、网络身份标识信息类中的“用户 ID、即时通信账号、网络社交用户账号、用户头像、昵称”等，这也进一步说明了需要引起足够重视。

(2) 敏感个人信息在标准中有部分提示，具体范围需要参考

敏感个人信息国家标准（此处应该是指《信息安全技术 敏感个人信息处理安全要求》（征求意见稿），现阶段可以依据 GB/T 35273—2020 进行判定），但存在提示不完全的情况，如个人交易信息、个人资产信息、个人借贷信息等。需要注意的是，标准中也强调了“特定身份信息”属于敏感个人信息，这承接了《个人信息保护法》的法规要求。

(3) 数据分类过程中业务视角尤为重要如个人信息中典型数据示例“兴趣爱好”，既在个人基本资料类也在个人标签信息类出现，因其有两种不同业务属性，同一个数据项可以划分在不同的分类中，这在金融行业表现得特别明显。这也是在实际操作中需要业务部门承担起本部门本领域数据分类分级工作的主要原因。

3.6 提供了重要数据识别流程、条件以及指南

从“表 1：数据级别确定规则表”来看，是否为重要数据的触发条件主要与影响对象和影响程度直接相关，以及行业领域主管(监管）部门评估确定。本文将简单归纳重要数据识别的流程，如图 2 所示。另外，《数据分类分级规则》附录 G 重要数据识别指南，给出了各行业、各领域需要考虑的 17 项因素，并列举了相应的示例，为各组织识别重要数据提供了详细的操作步骤。

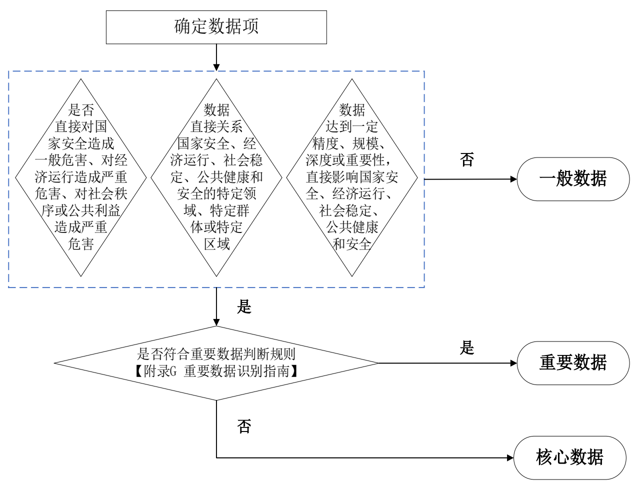


图 2 重要数据识别流程

在《数据分类分级规则》第 7 章中，要求行业领域主管（监管）部门“指导数据处理者准确识别、及时报送重要数据和核心数据目录信息”，数据处理者“对数据分类分级结果进行审核，形成数据分类分级清单、重要数据和核心数据目录，并对数据进行分类分级标识，按有关程序报送目录”。基于电信、金融、天津自贸试验区等行业、地区重要数据基本实践，本文整理了重要数据和核心数据报送目录，如图 3 所示，供大家参考。

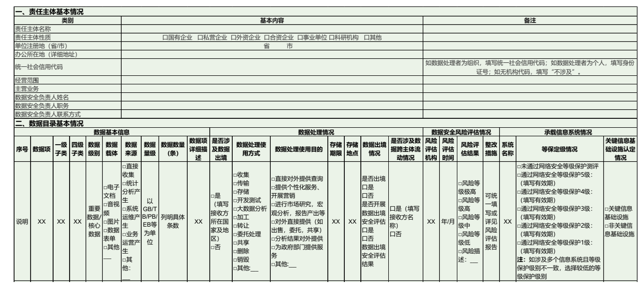


图 3 重要数据和核心数据报送目录

3.7 基于行业实践提供一般数据三种分级参考

《数据分类分级规则》附录 H“一般数据分级参考”，给出了一般数据三种分级参考，即一般数据可分为 2～4 个级别。关于组织适合哪种分级方式，本文结合行业要求及实践参考总结如下：

(1) 针对中、大型组织或业务特别复杂、数据类型特别丰富的组织，一般数据建议按照 4 个级别来进行划分。

(2) 针对中型组织或业务较为复杂、数据类型一般的组织，如连锁商超、宾馆、公园、景区等，一般数据建议按照 3 个级别来进行划分。

(3) 针对小型组织或业务单一、数据类型简单的组织，如个人商铺（包括但不限于街边小店、线上店铺），一般数据建议按照 2 个级别来进行划分。

(4) 行业有数据分类分级相关要求或标准指引的，如电信、工业、金融等行业，需要参照有关规定执行。

同时，也给出了特定类型数据的最低参考级别，如图 4 所示，具有很强的实践意义。

一般数据分级框架	特定类型一般数据的最低参考级别（不低于）	
一般数据分4级框架下	敏感个人信息	4级
	一般个人信息	2级
	组织内部员工个人信息	2级
	去标识化的个人信息	2级
	个人标签信息	2级
	有条件开放/共享的公共数据	2级
	禁止开放/共享的公共数据	4级
一般数据分3级框架下	敏感个人信息	3级
	一般个人信息	2级
	有条件开放/共享的公共数据	2级
	禁止开放/共享的公共数据	3级
一般数据分2级框架下	敏感个人信息	2级
	禁止开放/共享的公共数据	2级

图 4 特定类型数据的最低参考级别参考

4. 还存在哪些疑问

《数据分类分级规则》虽然在国家层面给出了具体的原则、框架、方法和流程，可指导行业主管部门制定本行业本领域数据分

类分级标准，但就组织开展具体分类分级工作而言仍缺乏许多细节。具体表现如下：

(1) 数据分级方法中，增加了“分级要素识别”这一重要步骤，其中要素的识别怎么与数据影响分析进行关联呢？虽然在级别确定规则步骤中，有四条规则进行了描述，但这不足以有效指导具体实践，仍存在诸多挑战。

(2) 由于业务的多样性，各行业数据分类分级会存在很大差异，组织可参考本标准进行业务属性分类，但由于每个人的理解和认知不同，数据分类分级结果存在较大差异性。

(3) 不同业务适配到同一个数据项时，可能存在数据级别不一样的情况。此时需要根据数据分级方法中多重因素进行级别确认。就如前面提到的典型数据示例“兴趣爱好”，在个人基本资料类时可能是 2 级，在个人标签信息类时可能是 3 级。

(4) 一般数据与重要数据之间的判定仍存在很多不确定性。如前所述，一般数据在一定条件下可能会被认定为重要数据，那么这个认定过程仍需要主管部门或相关部门来确认。

5. 总结

本文虽无法囊括整个数据分类分级具体内容，但可通过以小见大、见微知著的方式，为各位读者提供不一样的数据分类分级思路，相信大家一定能更加深入理解《数据分类分级规则》相关内容。最后，通过一个简单的示例将一般数据、重要数据、核心数据等做一个关联展示。

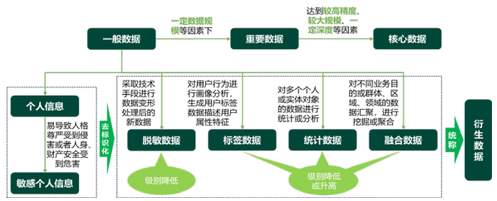


图 5 数据关联关系示例

针对已有行业标准的行业，如电信、金融、证券期货等，相关组织已经参照了行业标准正在或准备实施数据分类分级工作。那么，在应对国标《数据分类分级规则》的规则要求时，要注意几下几点：

(1) 数据分类方面，国标和行标最大不同点可能在于数据分类方法的细粒度上，国标在相关因素上考虑得比较详尽，都遵循了“业务条线—关键业务—业务属性分类”这一基础方式，可谓大家的共识。组织可在国标的基础上，优化自身的数据分类框架。

(2) 数据分级方面，国标以“数据分级要素”作为第一步，组织在原有行业标准的方法中，可根据自身特点选择对应定性（领域、群体、区域、重要性）、定量（精度、规模、覆盖度）和其他（深度）描述的分级要素调整自己的分级方法和步骤。需要注意的是，在具体实践中，不能简单粗暴地分为核心、重要、一般三个等级，需要对一般数据的等级再进一步细分，以 2 ～ 4 个细分等级为宜。

(3) 个人信息方面，在前面也讲到了，组织应在原有数据分类分级框架下，根据自身业务特点已有个人信息分类进行扩展，并结合数据分级方法确定对应数据等级。

针对未有行业标准的行业，相关组织在开展数据分类分级工作时，则需要参照国标《数据分类分级规则》，根据实际情况，对国标中相关内容进行裁剪，建立符合组织现状的数据分类分级规则，并动态开展数据分类分级工作。

以旧换新

秒变高配 旧设备变身新神器

“绿盟科技助你轻松升级

哪些能换？

- 使用时间超过八年及无法维保的硬件设备
- 硬件标准产品均可换，高、中、低配型号不设限，按需选配
- 支持旧设备与信创型号替换，助力落实国产化改造

有啥优惠？

- 换新惊喜折扣，直接让利
- 赠送维保服务：加送6个月维保
- 赠送安全情报通告服务：1年
- 云安全服务免费体验：

可选项目	体验时长/方式
网站安全监测	1个月
威胁检测与响应服务	1个月
外部攻击面管理 （含互联网资产核查）	1次
轻量化渗透测试服务	1次

旧的咋办？

- 参与换新活动后，旧设备不再享受技术支持及维保服务
- 符合条件的旧设备无需返厂

怎么换？

- 提供旧设备序列号及历史项目信息，联系您的销售经理或拨打4008186868
- 符合换新条件的旧设备均可参与活动，以下清单举例（包含但不限于），供您参考：

绿盟远程安全评估系统-RSAS

旧设备系列	参考可换系列
RSAS-S、X、E RSAS-S-CEC	RSASNX3-HXXX RSASNX3-EXXX

绿盟WEB应用防护系统-WAF

旧设备系列	参考可换系列
WAF-P1-6XX WAF-P1XXXX WAF-P16XXX WAF-P20XXX WAFNX3-P1XXX--6XXX	WAFNX3-HFAXX WAFNX5-HFAXX WAFNX3-HHXXX WAFNX5-HHXXX

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