cloud security alliance

CSA

CISCO

出版物

打造就绪网络，迎接“AI 风暴”

筑牢网络安全屏障，抵御利用 AI 提速的攻击

致谢

作者

Rich Mogull

平面设计

Stephen Smith

发起方简介

思科（纳斯达克股票代码：CSCO）作为全球技术领导者，正在全力重塑企业和组织在人工智能 (AI) 时代的互联与安全防护模式。40 多年来，思科一直致力于为世界提供安全可靠的连接。凭借业界领先的 AI 解决方案与服务，思科赋能客户、合作伙伴及社区激发创新潜能，提升生产力并筑牢数字韧性。

www.cisco.com/site/cn/zh/solutions/artificial-intelligence/security-in-ai-era.html

目录

简介

AI 模型的不断进步正在重塑网络安全攻防双方的关系，动摇了过去用于制定安全计划的传统风险假设。随着 Mythos 和 ChatGPT 5.5 等具备高级能力的模型陆续问世（详见《AI 漏洞风暴》(AI Vulnerability Storm)），AI 模型正快速迭代，在发现新型漏洞、编写全新漏洞利用程序以及执行复杂的多阶段自动化攻击方面的能力日益增强。过去几年的发展告诉我们，当今前沿模型的能力很快就会在广泛可用的基础模型上得到复刻，最终也能用于本地开源权重模型。

*我们过去用于制定安全计划的基础风险模型，如今已不再适用。*过去，发现新漏洞并编写新的漏洞利用程序需要数月甚至数年，而现在只需数小时。过去，发现这些漏洞并将其用于网络攻击需要深厚的技术能力，但现在这些门槛正在迅速瓦解。网络防御的经济效益正在发生改变；随着攻击所需的时间和技能门槛不断下降，我们以往的防御时间线已不再奏效，我们需要重新调整安全计划的方向。

虽然 AI 为攻防双方都提供了助力，但这种助力并不对等。正如《核心崩塌》(Core Collapse) 一文所述，这里存在数学层面的一个根本性失衡问题：攻击者可以将全部精力集中于单一问题（例如在*此处*寻找漏洞、为*此*开发漏洞利用程序，或攻击*那个*目标）；而防御者面临的则是组合复杂性难题，他们必须持续保护所有资源，抵御所有攻击者以及所有潜在的漏洞利用程序，以免功亏一篑。

只需观察单个新披露的漏洞所带来的影响，就能清晰洞察这种失衡。攻击者能够利用 AI 快速开发出漏洞利用工具，随即发起自动化攻击。而防御者即便拿到了补丁，也仍需识别出所有待修补的资源，在完成测试后，还需在不影响现有应用运行的前提下完成部署。无论我们多么优秀，这始终是一场遵循不同物理法则的竞赛。

不过，现有的某些策略有助于重新平衡这场博弈。从长远来看，随着所有新软件在发布前都利用 LLM 进行评估与加固，我们可以预期软件中漏洞的总量将会下降。而作为当下的防御者，我们还可以通过增设安全边界，*大幅提高成功实施攻击的组合复杂性*。

构建针对 AI 加持的攻击者的韧性，涉及各方面技术，这远远超出了安全团队自身能够独立管理的范畴。因此，我们需要全面强化各个领域的安全屏障，覆盖身份、工作负载、容器、API 和终端。

网络是我们强制执行这些边界的最重要工具之一，然而对于生产环境中运行的大多数现实网络而言，其现有的网络安全功能、设计逻辑及运营模式往往跟不上形势，防护能力严重不足。尽管多年来不断进步和投入，我们仍然拥有太多扁平化的网络，往往依赖过时的硬件，并且其网络运营和安全流程尚未为快速响应和自动化做好准备。在拥有复杂网络的企业和组织中，哪怕是更改一条防火墙规则这样简单的事情，也可能耗费数周甚至数月。

为了提高攻击复杂度（及成本），我们的网络应当：

- 迅速加快补丁更新频率，并确保持续进行针对性修补。
- 通过改进网络分段和限制高价值数据暴露的位置，缩小攻击的影响范围。
- 增加多个*串行（而非并行）*防御层，迫使攻击者绕过多个边界。例如，在单个应用栈的不同层级部署多个防火墙。
- 以响应迅速的运营流程为支撑，确保紧随威胁环境的变化而快速适应。

利用先进的 AI 能力进行防御，终将使精通 AI 的企业和组织在响应速度上具备足够竞争力，从而有效抵御攻击者快速开发漏洞利用工具带来的威胁。与此同时，安全边界将成为延缓攻击节奏的关键手段，为我们争取更新部署的宝贵时间。

本文旨在为网络及网络安全管理员提供指导。仅靠网络防御无法完全满足我们新的恢复能力需求，但增加网络安全边界是最有效的切入点之一。

打造现代化网络，实现敏捷防御

十多年来，整个行业一直在推动网络现代化，从改善边界防御到加强网络分段，近年来这一进程进一步加速。企业和组织上云之后，防御表现更加出色，因为云网络默认情况下是软件定义的，而软件定义网络有助于更轻松地创建和更改网络边界和网络分段。

但企业和组织在大规模落地网络现代化升级的过程中，始终面临客观存在的切实挑战。许多（如果不是大多数）物理网络，其内部结构仍然相对扁平。即使是高度灵活、易于分段且始终为软件定义的云网络，防护能力也未必更优越，根源大多在于企业和组织采用平移上云迁移方案，直接照搬老旧数据中心的设计。在实施零信任方案时，往往更多地关注员工和园区边缘的安全，而忽视了数据中心和应用间（东西向）流量的保护。企业和组织不一定自主持有并运维网络，而在第三方网络环境下，可观测性也会存在诸多限制。

许多企业和组织虽已开始实施微分段，但通常仅局限于其网络的一小部分（主要受限于实施的复杂性）。硬件大多较为陈旧，更新操作难度大、耗时长。此外，运营流程历来严谨有余而灵活性不足：以往除网络边界外，企业和组织从未真正迫切需要快速修补内部网络，因此这些流程主要侧重于边缘，导致内部网络修补的节奏大幅放缓。

这些做法本身并无差错，是针对当时所面临风险的合理举措。本文给出的大部分建议都是业内公认的成熟方案，许多企业和组织正在实施落地。问题在于，AI 改变了风险格局和防御时间窗口。不仅是网络边界，

我们的内部网络屏障也成为了安全边界的关键延伸，这对于遏制攻击者的 AI 优势至关重要。扁平化的网络意味着一旦有设备被攻破，其影响范围会更大。现在，漏洞利用程序的开发在数小时内即可完成，按部就班的缓慢变更将无法跟上其步伐。而且，即使是位于防火墙之后，那些无法快速修补的硬件也会变成更大的隐患。

好消息是，我们已经知道该怎么做，但新的挑战在于执行。我们需要快速行动，因为我们过去用来证明历史优先事项和时间安排合理性的底层风险假设已不再适用。我们在分段、可修补性和敏捷运营方面的网络安全基本原则在平衡局势方面可能非常有效，但现在，我们需要以更精细的方式、按照优先顺序、在紧迫的时间内执行任务。

网络工程/运营和网络安全团队也需要携手合作，共同解决这些问题。这些解决方案跨越了孤岛，需要同时具备基础架构、补丁部署和漏洞管理，以及 NGFW 和 WAF 等特定安全防御功能。本文的其余部分将阐述现代化的要素，以及如何分阶段实现这一目标。

基础要素与落地流程

为应对对抗性 AI 风险，网络现代化包含三个维度：

- **运营模式/流程：** 早期网络依靠物理布线连接硬件设备，防护体系围绕极少变更的静态 IP 搭建，这使得网络与网络安全流程普遍刻板繁琐，硬件升级、架构调整相关流程尤甚。这些流程是为依赖物理硬件的老旧网络环境打造的。补丁部署集中在边缘，内部硬件由于风险较低而处于次要地位。大多数网络实施了一层或两层防御，侧重于南北向流量以阻断来自互联网的威胁。云环境动态性更强，但不同场景差异悬殊，具体取决于成熟度、对直接迁移的依赖程度以及混合环境的依赖关系。安全模式已从*阶段性专项项目转变为常态化持续运维*。不仅漏洞修补、设备固件更新需要采取这种模式，网络分段工作也同样如此，需要在搭建架构的过程中增设并维护安全边界，而非一次性搭建架构之后便不再调整。可以将其视为两条并行运转的闭环流程：
    - **保持网络及时修补且处于最新状态**：目标是保持持续修补，并按照轻重缓急推进设备更新，确保无需被迫防护无法修复的网络资产。一直以来，这对网络设备而言极具挑战性，正因如此，我们建议将其作为一项全新的运营流程，并优先从面向外部的资源着手。这可以并且应该整合到您的 *VulnOps* 计划中。《漏洞风暴》报告中有更详细的描述：VulnOps 将漏洞管理视为一种具备持续自动化能力的 DevOps 式计划。
    - **更新分段与隔离**：随着应用的变更，向微分段演进，持续添加并维持多重边界（南北向及东西向），并结合采用拓扑与安全防御手段加以实施。

- **基础设施：**网络本身就是一个攻击目标，我们预计针对用于自动化攻击的网络硬件和软件的漏洞及利用程序的数量将不断增加。这如今已在发生；在我们撰写本文的前几周，已经发生了多次针对网络基础设施的攻击活动。因此，网络应当：
    - **保持最新状态**：及时更新硬件，确保其未停止生命周期且仍然受支持。
    - **可修补：**不仅更新硬件，而且集成到一个能够在数小时内（而非数周内）完成更新的补丁管理程序中。
    - **软件定义：**正如在云端所充分展示的，软件定义网络 (SDN) 更加敏捷，并且通常默认拒绝操作（具体取决于平台，例如软件定义边界 (SDP)），其中网络分段是基础状态，而非可选的附加措施。
- **架构：**架构定义了增加攻击者复杂性和成本的边界。我们可能无法立即在每个层级为所有资产都部署防火墙，但我们需要基于以下模块进行构建：
    - **边界安全：**这是我们目前最擅长的领域，而且这种需求不会消失。但是，大家将看到，我们需要对此进行调整，以适应更加持续的运营模式。在某些情况下，我们可能希望提供更多样化的平台并加强边界分段。
    - **宏分段：**对业务单元、应用栈和运营环境进行更好的分段。目标是将影响范围限制在网络内的单个 “栈” 中。
    - **微分段：**在应用栈的各层之间增设边界，同时也增加应用和服务之间的边界。
    - **零信任：**将这些边界与经过持续验证的身份和上下文感知策略相结合，确保按请求授予访问权限，而不是按网络位置。信任不再仅取决于 IP 或 VLAN，而是变得动态、遵循最小权限原则，并与身份和风险挂钩，而非网络拓扑。这包括软件定义边界 (SDP) 等技术。

支撑这一切的是可视性（我们拥有哪些资源，其位置及配置方式）、归属权（谁拥有什么资源、谁负责业务和技术决策），以及理想情况下的实时遥测数据。*我们并非在淡化实现这些建议的复杂性*，但现代技术确实使其更易实现，特别是针对单个项目而言。

我们还可以自己利用 AI 工具来进行持续测试与改进。将您自己的 LLM 指向您的网络，为它们提供（安全的）评估工具，并利用这些结果从攻击者的视角来审视自身的优势与劣势。

这一切都需要时间，有时甚至长达数年，而 AI 驱动的对抗性攻击可能已经开始。我们建议采取一种分阶段、按优先级、基于风险的方法，这样既能切合实际地快速取得阶段性成果，又能为实现长期目标奠定坚实的基础。

重要提醒：作为持续运营流程的一部分，以下阶段可并行进行。您无需完全完成一个阶段后再进入下一个阶段。

第 1 阶段：确定优先级，开展针对性的更新与加固

攻击者会优先渗透他们能触及的目标；而作为防御者，我们需要重点防护遭到攻陷后会带来最大损失的关键资产。因此第一步是，企业和组织需要识别外部受攻击面，并将其与自身优先事项保持一致。

在这个阶段，要秉持“由外而内，关键资产优先”的思路。首先，要确定面向互联网的受攻击面，同时也要识别出最重要的数字资产（即“皇冠上的明珠”）。在确定受攻击面时，我们指的是 *一切事物*。每一台设备和每一项服务，*包括您所有的安全和网络设备及服务*。任何接触互联网且可能从外部访问的一切事物。不要忘记 VPN 服务器、远程办公室、边缘路由器和基于云的虚拟安全设备等资产。

关键数字资产，即“保护面”，是指那些如果在数据泄露或破坏性攻击中受损，会对业务造成最大危害的应用和服务。针对这些服务，您需要制定一份完整的网络连接和数据流图谱。这些通常是网络和安全团队自身无法回答的业务问题。

然后进行交叉关联分析。针对给定的关键资产，从外部向内看，有哪些网络设备处于攻击者的访问线路上。防火墙、WAF、路由器、交换机、VPN（这可能不明显）……全都包括在内。使用以下图表来确定您的下一步行动:

状态	建议的操作
设备生命周期终止 (EoL)	更新硬件：淘汰和更换
设备处于扩展支持状态	尽快更换
补丁级别落后于当前版本	立即修补
设备未以高可用性 (HA) 方式部署（物理或虚拟设备）	增加 HA 对，以支持无停机修补
因流程或软件/硬件限制，设备无法在补丁发布当日完成修补	纳入 VulnOps 计划，以支持在发布当日完成修补，或更换为支持该功能的硬件

针对上述所有设备，还可以采取一些战术性措施。再次强调，我们深知这些措施实施起来并不简单，但它们是非常有价值的防御手段，能有效防范多种常用攻击技术:

- 至少应要求对管理操作启用 MFA，且理想情况下，管理员权限应通过 PAM 进行代理管控。
- 禁用那些暴露在互联网或 DMZ 上的管理接口。这里的管理接口涵盖对设备、SDN 控制器及其他管理工作站的管理访问入口。系统应拒绝来自不可信网络的所有流量。

- 启用出口过滤。这有助于减少攻击者命令与控制回调。请记住，这应包括基于 DNS 的过滤，而不仅仅是 IP。
- 如果您使用任何涉及网络/安全管理的自动化脚本或代码，请评估其中是否存储了密钥，并使用 LLM 来评估其是否存在安全缺陷。
- 为所有面向互联网的应用（设备、虚拟设备或云服务）部署 WAF。

第一阶段侧重于加固我们现有的资产，并确保其保持最新状态。我们需移除那些本身最有可能被攻破的高风险网络和安全工具。如果安全边界变成了攻击的入口点，那么它将毫无价值。

双重防护：保护 OT 和高关键性网络

运营技术 (OT)/工业控制系统可能很难或根本无法以任何常规频率进行修补（如果还能修补的话）。虽然 OT 过去通常存在于隔离环境中，但几十年来，我们越来越多地将其连接到 IT 网络，从而暴露了来自互联网的攻击路径。它们是首要目标和关键基础设施，却难以抵御 AI 驱动的漏洞、漏洞利用程序和攻击的快节奏威胁。

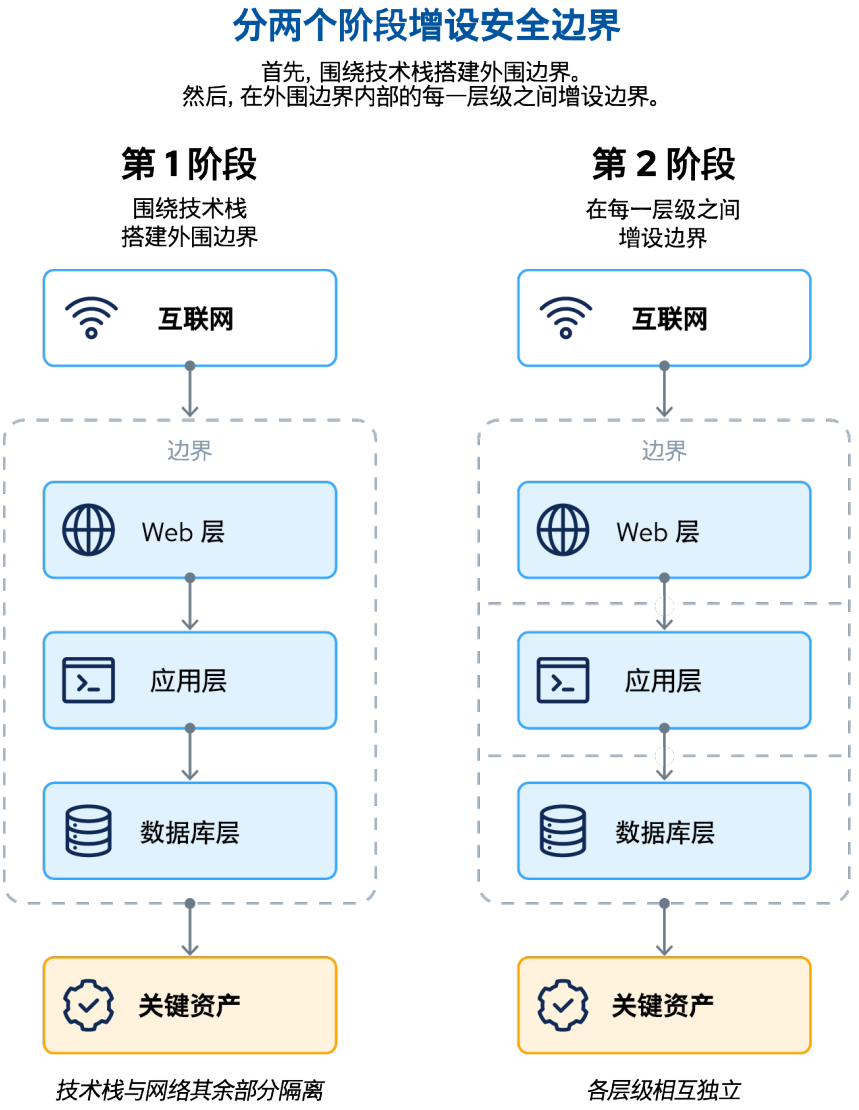
这些网络应当断开与互联网的任何路由连接。如果无法实现，建议采用类似数据二极管的单向传输技术来收集遥测数据。若这也难以实现，则应部署来自不同供应商的多台（2 台或以上）内联式下一代防火墙，并确保其实时更新。这样一来，攻击者需要多次利用零日漏洞才有可能渗透进 OT 网络。

第 2 阶段：分级管控，精准增设边界

第 1 阶段专注于加固已知资产，并淘汰那些无法修补或跟不上 AI 驱动补丁速度的过时硬件和软件。第 2 阶段将开始增加安全边界，以提升攻击者的组合攻击复杂度，并缩小成功攻击造成的影响范围。

第一步本质上是实施宏分段：即隔离技术栈，防止同一网络上那些价值较低、防御较弱的资源被攻陷后，成为攻击者轻易通往优先保护技术栈的跳板。在此环节有多种网络技术可以使用，包括部署防火墙、隔离子网和路由、迁移至隔离的虚拟网络、利用 SDN 控制策略进行隔离，或者直接跨越式地实施零信任。首要任务是防御东西向攻击，前提是您在第一阶段已经针对来自北方（互联网）的攻击完成了加固。

如果选择通过主机内部的工具来实现隔离，该工具必须具备底层运行能力，最好部署在操作系统之外（例如基于虚拟机监控程序），或者使用受外部管控的加固型代理程序。这样，即便攻击者成功入侵了主机，也无法轻易关闭网络安全防护。



第二步，也是更耗时且可能需要引入新技术的一步，是增加额外的*内联边界*。我们之前的步骤专注于减少容易利用的路径，但这一步增加了额外的、真实的攻击者成本。应用的每一层应仅与上一层和下一层通信，并且仅使用预期的端口和协议。SDN 是最佳选择，而且这种模式通常在云和完全虚拟化环境中更容易实施。

第 3 阶段：全域拓展，精细优化

第 1 和第 2 阶段的重点是，快速守护我们最重要的资产。我们淘汰过时的硬件与软件，将所有内容更新至最新状态，与防御较弱的资源进行隔离，然后开始添加安全边界，以增加攻击者的成本和复杂性。请记住，这仅仅是网络和网络安全层面的工作；正如我们的《AI 漏洞风暴就绪指南》所强调的，其他并行的安全举措同样至关重要。

第 3 阶段旨在完善这些防御措施并将其扩展至我们网络的其余部分。理想情况下，我们会从扁平网络演进到宏分段，再到微分段，最终实施零信任。尽管直接转向零信任是最理想的，但这在短期内可能难以实现。在宏分段方面，我们会着手将业务单元和信任区域进行隔离。

微分段将边界细化至单个工作负载和服务，确保每个单元仅与其实际需要的对象进行通信。如果大规模实施，不仅过程缓慢且成本高昂，因此通常需要按项目逐个实施。而这往往耗时数年，需要进行真实的流量和依赖关系映射才能执行任何变更，若操之过急，可能会导致生产环境瘫痪。

在 CSA，我们将微分段定义为零信任的一部分，但在一些网络和安全管理员看来，它们虽有区别，却在一定程度上存在重叠。微分段可以仅基于网络进行实施，并按照单个资产和单个应用实现隔离，而零信任则在连接决策中引入了身份和上下文。

零信任方案集所有这些功能于一身。一旦确定了精细的边界，并且您的策略跟随工作负载和身份而非网络位置，那么就可以针对每个连接单独授权访问。正如在云端已经证明的那样，这对攻击者来说是一个巨大的障碍。在实施过程中请留意您的瓶颈点：身份提供程序和其他整合的控制点会成为最高价值的攻击目标。正是因为所有系统都无条件信任它们，我们才必须对其施加更严格的审查、添加更多的边界，绝不能掉以轻心。

在这个阶段，绝对关键的是要严格审视技术管理流程和接口。毕竟，网络除了修补漏洞外还需要频繁更新，这就使得那些管理接口成为了攻击者垂涎欲滴的目标。

这一切永远没有真正的“完成”，这正是关键所在。能在 AI 驱动的攻击下屹立不倒的网络，不能止步于完成阶段性任务，而应将分段、修补和可视性常态化。这确保每个新应用上线即自带边界，让每次漏洞利用都面临攻击成本远高于防守成本的网络结构。安全事件仍会发生，但提升网络遥测能力也可以为事件响应团队提供强力支撑。

艰难转型，前景可期

本文涉及的技术并非首创，但其实施的规模、范围和时间规划，却与我们以往的运营模式截然不同。文中探讨的皆是基础性的安全原则，其中许多已在各种环境中久经实战考验。然而，作为网络与安全专业人士，我们必须始终平衡安全风险与管控措施给业务带来的阻力。增加安全控制措施会增加成本和复杂性，并可能拖慢部分业务进程。我们提出的方案实施难度高、而且肯定需要相应的投入，但现阶段没有更好的替代选择。

AI 改变了我们制定现有安全措施时所依据的基本风险假设，让大量技能较低的攻击者也能具备接近精英级漏洞挖掘者和渗透测试人员的攻击能力。我们再也不能指望拥有数周甚至数月时间来进行边缘修补，或者拥有数年时间来修补内部基础设施。我们也不能想当然地认为攻击者无法在我们复杂的内部架构中穿行，更不能心存侥幸，觉得只有一家分支机构在使用、未打补丁的 VPN 服务器不会引来攻击者，因而疏于防护。

尽管从目前攻防博弈的客观态势来看，进攻方占据优势，但实现攻防平衡的路径清晰可循。最终，随着我们将同样的 AI 能力全面融入到软件开发与安全评估中，所有软件在发布之初就会具备原生安全能力。与此同时，我们应推进网络现代化并落实零信任架构，增加攻击者需要绕过的安全边界数量，从而抬升他们实施攻击的综合成本与复杂度。当我们能够持续进行修补和更新之后，留给攻击者实施攻击的窗口期将大幅压缩。

只有这种多重举措协同发力的综合策略才能扭转攻防失衡的局面，让我们占据优势，而且哪怕只是做出微小的调整，也能带来巨大的收益。升级运维模式，提升响应能力并整合 VulnOps。优先专注于保护关键资产。确保全面升级软硬件，保障漏洞补丁发布当日即可完成修补。尽可能迁移到 SDN，包括大多数虚拟网络和云网络，因为 SDN 能够极大降低网络分段的落地与维护成本。从东西向流量分段入手，隔离各技术栈，再逐步增设内嵌式南北向安全边界，最终落地零信任架构与 SDP。

是主动落地零信任防护，还是被动应对零日漏洞攻击，您来抉择。

要想在 AI 时代蓬勃发展，企业和组织并非需要试图消除每个漏洞，而是需要始终领先于试图利用漏洞的攻击者，持续缩小可被利用的受攻击面。各类安全防护壁垒能够为防御者争取时间、抬高攻击者的攻击成本。缩减受攻击面、提升响应速度、增设多层安全边界三管齐下，方能构建真正的网络安全韧性。