



赋能当下创新， 引领未来突破

实用战略，助力打造
AI 就绪数据中心



当基础设施不堪重负之时

想必您对此深有体会：日复一日，关于性能峰值的告警接连不断。问题的症结在于，任务关键型应用仍在传统基础设施上运行，而这些基础设施已不堪重负，无法应对当今庞大的业务规模。停机下线、全面重构？这根本行不通。随着工作负载不断增加，现有的计算、网络和存储能力已濒临极限。您的团队大概率疲于应付突发问题，几乎没有余力进行前瞻性规划。

与此同时，业务端正全力向 AI 领域挺进，随之而来的计算与基础设施需求前所未有，令人措手不及。各方期望不断提升：既要推进现代化转型，又要加快行动步伐，还要保障安全，而且不能停机，也不能超支或增员。

当下，常规升级已无济于事，您需要打造全新数据中心，来为各类工作负载提供支持，随时灵活扩展以适应指数级增长，保障全栈安全并实现统一管理，并且始终保持强大韧性。领导者需要果断采取行动，引领企业和组织从容应对未知挑战，始终立于不败之地。

当前挑战:

突破基础设施极限， 全面实施转型

对于首席信息官 (CIO) 和高级 IT 主管而言，AI 时代意味着一种新的常态：机遇前所未有的，挑战同样不容小觑，比如成本压力巨大、对于快速交付 AI 能力与使用案例的业务需求愈发紧迫、威胁形势日益严峻。

- 尽管 AI 投入持续增加，但业界对于其总体成本仍缺乏清晰认知。Gartner 的分析¹指出，这类持续性支出可能飙升至预期成本的 5 到 10 倍，面临着极大的失控风险。
- 为赢得竞争优势，企业和组织竞相开发 AI 原生应用。因此，唯有不断压缩推理时间和训练周期，才能在竞争中保持领先。生成式 AI (GenAI)、大语言模型 (LLM) 以及智能体 AI 的兴起，进一步加快了 IT 领域的变革与创新步伐。
- 恶意攻击者利用 AI 肆意发动攻击；与此同时，AI 自身也引发了一系列全新的安全隐患与防护难题。

这对基础设施而言意味着怎样的考验？面对当前形势，传统数据中心早已力不从心。试想一下：如今，推理与训练工作负载的计算密度是传统应用的 20 倍，功耗更高达后者的 10 倍。² 面对 AI 产生的海量数据，现有的网络、计算和存储系统早已不堪重负。再加上边缘侧的实时处理需求，传统基础设施的种种缺陷已然暴露无遗，不容忽视。

归根结底：我们正处于一个关键的历史拐点。

AI 就绪的数据中心 应具备哪些条件？

1 无处不在：全面支持各类工作负载

传统工作负载与 AI 工作负载的并存已成必然。数据中心基础设施必须能够从容应对不断增加的传统工作负载和 AI 工作负载，不受其部署位置的制约。无论是在本地、云端还是边缘侧，都应 AI 部署在靠近数据的位置，而不是反其道而行之。

2 灵活扩展：从容应对指数级增长

企业和组织必须将高密度计算、无损低延迟网络以及可持续发展置于首要位置。提升能效已不再是锦上添花，而是不可回避的战略命题。

3 纵深防御：为全栈安全保驾护航

必须将安全防护深度融入数据中心架构，确保在基础设施、网络、应用乃至 AI 模型本身等各个层级贯彻执行。

4 化繁为简：大幅精简运营流程

对传统工作负载与 AI 工作负载实施统一管理，有助于团队顺畅开展运营，并以标准化的方式高效管理所有工作负载。

5 内置韧性：始终保障业务连续

面对突发中断事件，企业和组织必须迅速完成检测、响应与修复。

思科： 全面赋能 现代数据 中心





赋能数据中心的 关键要素



强大、灵活的
基础设施



焕新重构的安全防护



全面统一的可观测性
和管理



强大、灵活的 基础设施

应对 AI 浪潮，亟需打造新一代基础设施：必须具备大规模扩展、超低延迟以及与计算、网络和存储无缝集成的能力。思科提供高度灵活的部署选项：既可作为全栈系统进行整体部署，也可作为数据中心基础模块进行独立部署，恰好能完美满足这一需求。



基础模块：网络、芯片、 光模块与计算

- 借助 Nexus 交换矩阵，您可以构建自动化的高性能以太网架构，该架构不仅能支持传统工作负载，更能为前后端网络中要求严苛的 AI 工作负载提供强大的 AI 网络支撑。
- 其核心在于 Cisco Silicon One，这款业界领先的统一网络芯片架构纵贯网络边缘与核心，可全面驱动性能突破、大幅提升运营效率并简化管理。
- 无论是在 AI 的后端还是前端网络，思科光模块都可为交换机与服务器提供极具韧性且高度可靠的连接。
- 思科 Unified Compute System (UCS) 将计算、网络和存储紧密集成到同一敏捷平台之中，从根本上革新了数据中心架构。

除提供上述数据中心基础模块外，思科亦可助您依托经过全面集成和预验证的全栈系统打造 Secure AI Factory，该系统专为 AI 工作负载而设计，可大幅加速 AI 应用的开发与部署进程。

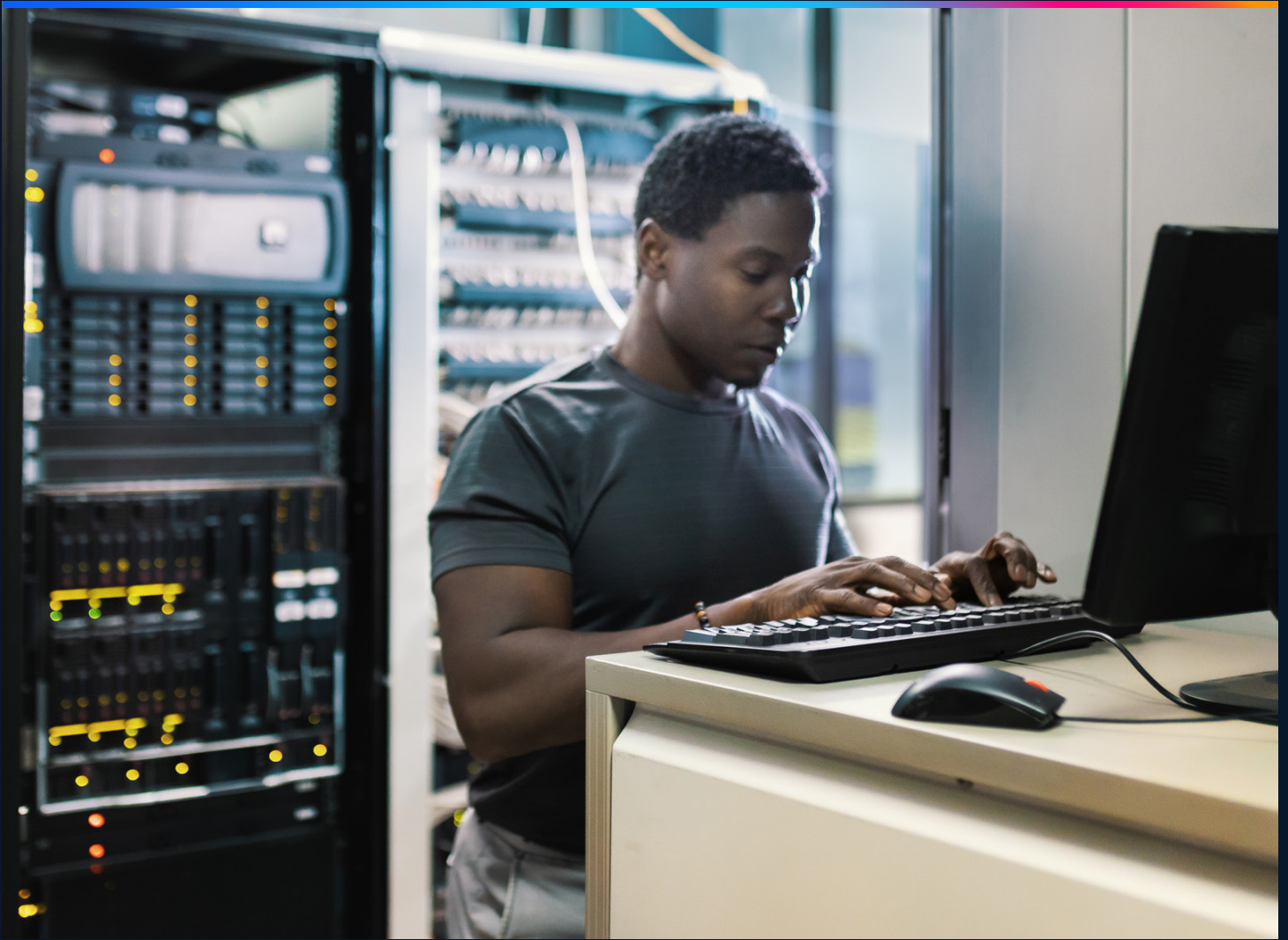
全栈系统

AI POD

Cisco AI POD 可简化 AI 部署，并根据业务需求灵活扩展。AI POD 集成了高性能的 UCS-X 计算节点、Red Hat OpenShift（用于 Kubernetes 协调）以及 NVIDIA AI Enterprise（专为优化 AI 工作流程打造），同时在全栈架构中融入思科安全防护，旨在全方位保护数据、模型和基础设施。这种一体化方法有助于简化部署、加速 AI 的落地应用并确保交付始终一致的成果。

融合基础设施

我们的融合基础设施同样基于思科验证设计构建，将日立、VAST Data 和 NetApp 等合作伙伴的前沿存储技术与思科计算及网络集成，打造出无缝的端到端解决方案。



焕新重构的安全防护

随着工作负载在混合环境中变得日益分散，一致地执行安全策略变得愈发困难。我们的安全产品组合不仅覆盖全面，而且实现了深度集成，旨在为您提供更精准的威胁检测、高度一致的策略执行以及简便的管理。借助 Cisco Hybrid Mesh Firewall，您可以在全栈（涵盖基础设施、工作负载、AI 模型等）各个层级内置执行点，实现从本地环境到云端的全方位安全防护。此外，借助 Security Cloud Control，您能够轻松管理全栈的安全策略，只需定义一次策略，即可在所有位置统一执行。

思科还通过 Hypershield 与 AI Defense，为数据中心安全领域带来独树一帜的创新突破。Hypershield 可为数据中心提供 AI 原生的安全防护，并在其内部构建起一张由分布式执行点交织而成的安全网。它在工作负载、网络和基础设施等层级均可提供深度可视性和精细化的策略执行。AI Defense 开创了一种全新的防护模式，在 AI 模型与 AI 应用的开发过程中同步提供安全保障。AI Defense 采用创新的算法红队机制，对 AI 模型进行持续验证，实时检测模型行为是否符合预期。当多数企业和组织还在耗费 7 至

10 周时间来验证模型时，AI Defense 已将该过程大幅压缩至几分钟之内。



统一的可观测性和管理

除了保障安全之外，思科还可提供覆盖整个环境的可视性和洞察，助您提高数字韧性。ThousandEyes 可助您全面洞察自有和非自有网络，确保无缝连接，这一点至关重要，因为在公开报告的中断事件中，大多数问题都源于非自有网络。Splunk 可观测性云则将这种可视性扩展到应用和基础设施，大幅提升性能问题的检测速度与解决效率。

借助 Nexus One 和 Intersight 的统一管理，IT 团队可轻松管理传统工作负载和 AI 工作负载。团队借此能够大幅简化计算与网络环境的运营，加快问题解决速度并提高可靠性。



赋能业务全局， 交付关键成果

如今，IT 主管必须能够加快行动步伐，推进 AI 的规模化应用，同时管理日益复杂的混合环境。思科独树一帜，将网络、计算、安全和可观测性深度整合，助力企业和组织满足这些需求。这种平台化方法能随着时间的推移产生复利效应，让每一笔新增投资都能实现价值倍增，同时让运营化繁为简。

思科可助力实现以下四项重要成果：

加快部署速度

凭借深厚的专业积淀、预验证设计以及专为严苛环境打造的关键基础设施，我们能够助您加快部署速度并缩短价值实现时间。

在快速增长与规模化扩张的双重压力下，Workday 数据中心容量趋近极限，网络亦难以满足发展需求。部署思科解决方案后，Workday 成功将网络部署耗时缩短 83%，一举突破了响应速度与扩展能力的双重瓶颈，为业务的快速扩张扫清了障碍。³ 资深首席工程师 Ahsan Ghayas 表示：“我们在传统数据中心使用 Cisco Nexus 交换机的体验非常棒。因此，在打造云架构网络时，Cisco Nexus 9000 系列交换机自然成为了我们的不二之选。该系列交换机在端口密度、容量和可扩展性方面表现出色，还具备我们非常看重的线速加密功能。”

推动 AI 规模化落地

无论是数据中心还是边缘侧，思科均可依托 Nexus 网络、Silicon One、光模块和 UCS 计算，提供 AI 所需的可扩展、高效基础设施。思科 IT 团队的亲身实践便是有力印证。面对产品团队的紧迫需求，思科 IT 团队必须在三个月内搭建一套 AI 集群，针对思科广泛的产品组合开发和测试全新 AI 功能，涵盖模型训练与推理等。最终，团队成功部署了 AI 就绪数据中心，其部署耗时比传统方式锐减了 80%。⁴

打造无处不在的安全防护

Cisco Hybrid Mesh Firewall 提供分布广泛的执行点，确保安全机制得以深度融合全栈之中。借助 Hypershield 等创新技术，思科成功将安全防护直接融入数据中心架构本身。例如，在思科智能交换机上运行 Hypershield，即可在交换机层面统管安全策略并实现内联网络分段，从而提升安全与运营效率，并减少所需管理的硬件数量。通过部署 Cisco Secure Workload (Cisco Hybrid Mesh Firewall 的关键组件)，美联航得以深入洞察其混合多云环境，实现自动化的威胁响应，并实施微分段。⁵

实现全面可视性

通过将 ThousandEyes 与 Splunk 的卓越能力集成至整个技术栈，企业和组织能够全面洞察其公共云和私有云、自有网络和第三方网络以及数字服务。从而更快检测并解决问题、优化性能，并始终对整个环境了如指掌。

例如，信达金融公司借助 Cisco ThousandEyes 和 Splunk 获得了深度可视性，加快了威胁检测速度，并简化了事件响应流程。

信达金融公司网络基础设施副总裁 Jermaine Mason 表示：“ThousandEyes 让我们能够直观了解那些日常依赖的非自有网络的状况。我们得以在客户感知到问题之前，主动上报并介入处理。”⁶

在这些优势的加持下，您可以在成本效益、能效和可持续发展方面获得更丰厚的回报。思科产品组合专为提高效率与推动可持续发展而打造，其网络与计算解决方案能够降低能耗、减少占用空间，并延长产品生命周期。无论是荣膺 SEAL 奖“年度可持续产品”称号的 UCS X 系列，还是入选 2025 年 Gartner 数据中心交换产品魔力象限的 Nexus 9000 系列，均可通过高效的供电、冷却设计和实时能源洞察，助力企业和组织降低成本并实现可持续发展目标。

简单三步，实现 AI 就绪

1

焕新升级：

改造既有系统，为未来扩展奠基

即便 AI 尚未成为战略核心，推进现代化升级与前期准备工作亦是当务之急。战略性升级能够推动传统系统更新、保障混合及多云环境的安全、提升网络性能，并为未来的技术演进奠定坚实基础。重点应放在打造高度协同的集成化技术栈上。功耗和热效率同样不容忽视：这方面的早期设计抉择将在基础设施不断扩展以满足 AI 需求的过程中，大幅降低长期成本，实现化繁为简。

2

深度优化：

针对 AI 工作负载改善现有环境

在着手部署 AI 工作负载时，涵盖整个环境的可视性、韧性和集成安全功能显得尤为重要。优化的核心不在于单纯扩充算力，而在于打造针对 GPU 优化的模块化基础，轻松、安全实现扩展。AI 的引入必然伴随着全新的复杂性，如果无法全面洞察工作负载和基础设施，性能瓶颈与运营风险便会很快显现。这一阶段，应重点引入遥测机制，实现性能监控自动化，同时确保安全防护贯穿于从基础设施到 AI 模型各个层级。尽早夯实这一基础，可确保 AI 部署项目在不断扩展的过程中，始终保持出色的韧性、高效运转并受到严密保护。

3

全力提速：

在规模化扩展阶段从容应对

一旦 AI 从试点阶段迈入规模化生产，速度将成为决定竞争胜负的关键因素，基础设施必须具备即时响应的能力。CIO 应提前布局规模化扩展所需的能力，包括高性能 GPU 加速算力、低延迟网络架构以及高带宽连接，满足生成式 AI 和大型语言模型等高级使用案例的需求。在这一阶段，运营敏捷性的关键在于对 AI 生命周期的全面可视性，从工作负载调度、GPU 利用率，到模型性能与业务影响，每个环节都要清晰可见。凭借合适的基础设施，您可以在保持掌控力的前提下加快行动步伐，并赋能团队笃定创新，交付可衡量的 AI 价值。

资源

1. 思科: [“重塑全球网络架构”](#)
2. 思科: [“思科快速部署 AI 就绪数据中心, 全面满足未来扩展需求”](#)
3. 思科: [“美联航携手思科起航: 筑牢数字韧性基石”](#)
4. 思科: [“数据中心现代化, 走适合自己的路”](#)
5. 思科: [“领跑智能体 AI 时代”](#)
6. 思科: [“人工智能就绪指数”](#)
7. 思科: [“升级技术, 尽释潜能: 成功转型的三大业务驱动因素”](#)

参考资料

1. Gartner: [“AI 定价: 遏制成本激增的应对策略”](#)
2. CIO.com: [“AI 成本与复杂性危及企业价值, CIO 压力与日俱增”](#)
3. BzTech: [“直击 Cisco Live 2025 大会: 信达金融公司数据可视性提升之道”](#)

未来已至，时不我待。 即刻迈向 AI 就绪

步入 AI 时代，渐进式升级已然不足以满足需求，唯有大胆重塑数据中心战略方能制胜。借助思科的平台化方法，无论是当下还是未来，您都可以信心十足地推进现代化升级，安全进行扩展，并充分释放 AI 的价值。

[了解如何](#)重塑数据中心，迎接 AI 时代。

