

AI 就绪型企业级平台

通过 NVIDIA 和 VMware 共同提供的端到端平台，为每个企业释放 AI 的强大力量

企业级 AI 工作负载所面临的主要挑战

- **风险** - 很难从不同的产品中整合端到端 AI 解决方案，并将其与现有基础架构集成。
- **性能** - 对于 AI、机器学习和数据分析工作负载而言，拥有高性能具有至关重要的意义。快速部署也很关键。
- **扩展** - 从概念证明到企业级部署都需要高效的扩展（通过高效使用资源来确保可管理性和可用性）以及基础架构成本管理。

无论是改善金融服务领域的客户关系、精简制造商供应链，还是帮助医生为患者提供更好的治疗效果，人工智能 (AI) 都在改变每个行业。虽然大多数企业都明白，他们需要投资于 AI 来确保其未来发展，但他们很难找到助力取得成功的战略和平台。

与传统的企业级应用不同，AI 应用对于许多 IT 部门来说是相对较新的发展领域。他们依赖于快速发展的开源前沿代码，缺乏经过验证的方法来满足企业中规模化生产环境的严苛要求。事实上，Gartner 研究显示：“平均只有 54% 的 AI 项目能够从试点阶段进入生产阶段”。¹

NVIDIA 与 VMware 共同推出 AI 就绪型平台

VMware 与 NVIDIA 强强联手，通过提供针对 AI 工作负载进行了优化的端到端企业级平台，来为每个企业释放 AI 的强大力量。该集成式平台可提供一流的 AI 软件，即 NVIDIA AI Enterprise 套包，该软件针对企业级工作负载平台 VMware vSphere® 进行了优化并获得了相关认证。该平台在行业领先的加速服务器 NVIDIA-Certified Systems™ 上运行，可加快开发人员构建 AI 和高性能数据分析的速度，使企业能够在已投资的相同 vSphere 基础架构上扩展现代工作负载，并提供企业级可管理性、安全性和可用性。



为企业带来的优势

- **放心部署** - 利用经联合认证的 NVIDIA 和 VMware 解决方案，通过优化的软件部署和管理 AI 工作负载。
- **提升 AI 性能** - 为 AI 训练和推理提供近乎裸机的性能。
- **在不影响性能的情况下进行扩展** - 通过新的横向扩展功能，进行更为复杂的大规模 AI 训练和数据分析。

为企业带来的优势

- 通过利用虚拟化将 AI 部署整合到现有企业级基础架构中，避免出现 AI 孤立小环境并简化管理。
- 获得近乎裸机的 AI 和机器学习训练和推理性能，并获得企业级安全性、可管理性和效率等额外优势。
- 实现 AI 工作负载的高可用性，同时简化整合、扩展和升级等基础架构维护工作。
- 通过自动执行工作负载初始安置简化整个 AI 基础架构的大规模管理，以优化资源消耗和确定性成本。

关键促成技术

该解决方案由以下构造块组成。

NVIDIA AI Enterprise 套包

NVIDIA AI Enterprise 是一套端到端、云原生的 AI 和数据科学应用及框架，经过 NVIDIA 优化和独家认证，可通过 NVIDIA-Certified Systems 在 vSphere 上运行。它包括来自 NVIDIA 的关键促成技术和软件，可在现代混合云中快速部署、管理和扩展 AI 工作负载。NVIDIA AI Enterprise 由 NVIDIA 授权许可并提供支持。



vSphere

vSphere 是企业级工作负载平台，可在全球数十万个企业中运行超过 7,000 万个工作负载。vSphere 可将裸机服务器（包括基于 CPU 和 GPU 的资源）转换为集中管理的 AI 和机器学习基础架构池，从而能够按需快速置备虚拟机和容器。

vSphere 可优化性能、提高可用性、增强安全性、精简维护并降低成本，从而创建敏捷、高效、具有恢复能力且原生安全的基础架构平台，以支持现有工作负载和新一代应用（如 AI）。

vSphere 的功能已得到增强，可利用 NVIDIA 的最新 GPU 以最佳方式运行 NVIDIA AI Enterprise 软件套包。

NVIDIA-Certified Systems

NVIDIA AI Enterprise 套包经过认证，可在 NVIDIA-Certified Systems 上运行，该服务器包含：

- 基于 NVIDIA Ampere 体系架构的 GPU，例如 NVIDIA A100 Tensor Core GPU。Ampere 体系架构中包含的 Tensor Core 技术可显著加快 AI 运行速度，将训练时间从数周缩短到数小时，并极大地加快推理速度。
- NVIDIA® Mellanox® ConnectX® SmartNIC 和 NVIDIA BlueField® 数据处理单元 (DPU) 提供一系列软件定义的硬件引擎，以加速网络连接并提升安全性。这些组件可实现两全其美：一流的 AI 训练和推理性能，以及所有必要级别的企业数据隐私性、完整性和可靠性。

开始体验

要了解更多信息，请访问 vmware.com/aiml

要了解有关 VMware 和 NVIDIA 合作关系的更多信息，请访问 www.vmware.com/cn/partners/global-alliances/nvidia 和 www.nvidia.com/zh-cn/data-center/vmware/

或者联系销售部门，网址为：

www.vmware.com/cn/company/contact_sales 和 www.nvidia.com/zh-cn/contact/sales/。

1. Gartner®, “Press Release: Gartner Survey Reveals 80% of Executives Think Automation Can Be Applied to Any Business Decision”, 2022 年 8 月 22 日。GARTNER 是 Gartner, Inc. 和/或其附属机构在美国和其他国家/地区的注册商标和服务标志，经许可在此使用。保留所有权利