

Compatibilidad con los servidores Cisco AI UCS que contienen GPU Nvidia

Contenido

[Introducción](#)

[Antecedentes](#)

[Descripción general del software NVIDIA AI Enterprise](#)

[¿Qué es NVIDIA AI Enterprise?](#)

[Componentes clave](#)

[Modelos de implementación](#)

[Ventajas de la compatibilidad con la licencia empresarial NVIDIA AI](#)

[Información para efectuar pedidos](#)

[Cómo obtener asistencia](#)

[Cuándo ponerse en contacto con CiscoSupport](#)

[Cuándo ponerse en contacto con el soporte técnico de NVIDIA](#)

[Compatibilidad con otro software relacionado con la IA](#)

Introducción

En este documento se describe la suite NVIDIA AI Enterprise Software (NVAIE), así como las completas ventajas de asistencia técnica disponibles cuando se implementa en plataformas Cisco Unified Computing System (UCS). Este documento también describe cómo involucrar a la asistencia de Cisco y NVIDIA para una implementación óptima y asistencia operativa.

Antecedentes

A medida que la inteligencia artificial y las cargas de trabajo de aprendizaje automatizado se convierten en elementos fundamentales para las operaciones empresariales, las organizaciones necesitan no solo una infraestructura potente, sino también software y asistencia de clase empresarial. El software NVIDIA AI Enterprise ofrece una plataforma de IA completa, segura y compatible, optimizada para su implementación en la infraestructura de Cisco UCS.

Este documento describe:

- Qué conlleva el software NVIDIA AI Enterprise
- Ventajas de soporte de la licencia de la suite de software NAVE
- Cómo interactuar correctamente con los canales de asistencia de NAVE en Cisco UCS

Descripción general del software NVIDIA AI Enterprise

[¿Qué es NVIDIA AI Enterprise?](#)

NVIDIA AI Enterprise (NVAIE) es una completa suite de software que proporciona seguridad, estabilidad, capacidad de gestión y compatibilidad de clase empresarial para cargas de trabajo de IA que se ejecutan en entornos virtualizados y nativos de la nube. Está optimizado y certificado para su ejecución en las principales plataformas empresariales, incluido Cisco UCS.

Componentes clave



Nota: Para obtener una lista actual y exhaustiva, visite [NVIDIA AI Enterprise](#)

Marcos y herramientas de IA:

- Flujos de trabajo de NVIDIA AI: flujos de trabajo predefinidos de extremo a extremo para aplicaciones comunes de IA
- NVIDIA TAO Toolkit - Transferencia de herramientas de aprendizaje para la visión de la computadora, la inteligencia artificial conversacional, y más
- NVIDIA Riva - SDK para la creación de aplicaciones de inteligencia artificial de voz
- NVIDIA Morpheus: estructura de inteligencia artificial de ciberseguridad
- NVIDIA Metropolis: plataforma Vision AI para análisis de vídeo inteligente
- NVIDIA Clara - Marco de aplicación de la IA sanitaria

Software de infraestructura:

- Software NVIDIA vGPU: tecnología de virtualización de GPU (con licencia por separado o incluida con NVAE)
- NVIDIA GPU Operator - operador Kubernetes para la gestión de GPU
- Operador de red NVIDIA - Gestiona los recursos de red NVIDIA en Kubernetes
- NVIDIA Container Toolkit - Herramientas para crear y ejecutar contenedores acelerados por GPU

Pila de desarrollo de IA:

- Estructuras de IA optimizadas (TensorFlow, PyTorch, JAX, etc.)
- NVIDIA RAPIDS para la ciencia de los datos
- NVIDIA Triton Inference Server para la implementación de la IA de producción
- NVIDIA TensorRT para inferencia optimizada
- Bibliotecas NVIDIA cuDNN, cuBLAS y CUDA

Gestión y seguridad:

- NVIDIA Base Command Manager Essentials: gestión de infraestructura para IA
- NVIDIA License System: gestión centralizada de licencias
- Versiones de parches de seguridad y compatibilidad a largo plazo (LTS)

Modelos de implementación

1. Entornos virtualizados: VMware vSphere, virtualización de Red Hat, GPU virtual NVIDIA
2. Sin software específico: implementación directa en servidores físicos
3. Nativas de la nube: Kubernetes e implementaciones en contenedores
4. Nube híbrida y multinube: plataforma de IA coherente en todos los entornos

Ventajas de la compatibilidad con la licencia empresarial NVIDIA AI

Asistencia empresarial integral

Con la licencia NVIDIA AI Enterprise, obtendrá importantes ventajas de soporte técnico que van más allá del acceso básico al controlador de la GPU.

1. Asistencia para software de nivel de producción

- Actualizaciones y parches de software: actualizaciones de seguridad periódicas, correcciones de errores y mejoras de funciones
- Versiones de asistencia a largo plazo: versiones de LTS estables con períodos de asistencia ampliados (hasta 2 años)



Nota: Los niveles de asistencia de Nvidia (básico frente a las 24 horas del día, los 7 días de la semana) se basan en la licencia específica adquirida como se describe [aquí](#)

Componente de servicio	Estándar	24 x 7
Acceso a asistencia técnica	Horario comercial local	24 horas, 7 días a la semana
Tiempos de respuesta inicial de gravedad 1	4 horas laborables	1 Hora
Tiempos de respuesta inicial de gravedad 2	4 horas laborables	2 horas
Tiempos de respuesta inicial de gravedad 3	1 día laborable	4 horas laborables
Gravedad 4 Tiempos de respuesta inicial	2 días laborables	1 día laborable
Acceso al portal del cliente	24 x 7	24 x 7
Web	24 x 7	24 x 7
Teléfono	Horario comercial local	24 x 7

2. Asesoramiento técnico y orientación

- Asistencia en la implementación de IA - Ayuda con el diseño de arquitectura y las prácticas

recomendadas

- Optimización del rendimiento: orientación para ajustar las cargas de trabajo de IA para obtener el máximo rendimiento
- Soporte de Framework - Asistencia con TensorFlow, PyTorch y otros marcos de IA incluidos en el paquete
- Contenedor y Kubernetes Support - Ayuda con implementaciones de IA nativas de la nube
- Acceso a la base de conocimientos: documentación completa, guías y recursos para la resolución de problemas

3. Ventajas de licencias y acceso

- Software verificado y validado: versiones probadas y certificadas por la empresa
- Portabilidad de licencias: flexibilidad para trasladar licencias entre infraestructuras
- Administración centralizada de licencias: sistema de licencias NVIDIA para administrar las implementaciones
- Modelo de suscripción: opciones de licencia anual o plurianual
- Software Download Access - Portal oficial de descargas para todos los componentes de la NVAIE

4. Seguridad y cumplimiento

- Parches de seguridad: remediación oportuna de vulnerabilidades
- Seguimiento de CVE: notificación proactiva de problemas de seguridad
- Documentación de conformidad: compatibilidad con los requisitos normativos
- Secure Software Supply Chain - Componentes de software verificados y firmados

Cobertura del soporte

Función	NVIDIA AI Enterprise
Acceso al controlador de GPU	✓
Documentación básica	✓
Foros de comunidad	✓
Suite de software AI	✓
Asistencia empresarial (24 horas al día, 7 días a la semana)	✓
Tiempos de respuesta de SLA	✓

Función	NVIDIA AI Enterprise
Parches de seguridad	✓
Asistencia a largo plazo	✓
Certificación de producción	✓
Gestión técnica de cuentas	Complemento opcional
Acceso a flujo de trabajo AI	✓
Herramientas de gestión	✓

Información para efectuar pedidos

Cisco UCS con NVIDIA AI Enterprise

Paso 1: Seleccionar hardware de Cisco UCS

Elija la plataforma Cisco UCS y la configuración de GPU NVIDIA adecuadas:

Configuraciones de ejemplo:

- UCS C885A M8 con GPU NVIDIA HGX H200
- UCS C845A M8 con GPU NVIDIA RTXP6000
- UCS C880A M8 con GPU NVIDIA HGX B300
- UCS C240 M7 con GPU NVIDIA L40S
- UCS serie X con GPU NVIDIA H100 para implementaciones modulares



Nota: Cada NVIDIA H200 NVL, H100 PCIe y H100 NVL GPU incluye una suscripción NVIDIA Enterprise de cinco años, que incluye NVIDIA AI Enterprise. Es necesaria la activación del software. Si es necesario, debe activar la licencia y abrir un caso de soporte de software con NVIDIA.

Haga clic [aquí](#) para obtener más información.

Por ejemplo, UCS C845A M8, UCS X-210C-M8, UCS C240 M7 con GPU NVIDIA H200 incluyen una suscripción empresarial NVIDIA de cinco años.

Paso 2: Licencia Enterprise de NVIDIA AI

Las licencias NVIDIA AI Enterprise se venden por separado y se pueden adquirir a través de:

- Cisco Commerce: incluye compras de Cisco UCS



Nota: La licencia de NVIDIA también se puede adquirir fuera de Cisco.

Tipos de licencias:

- NVAIE- Estándar: suscripción anual estándar (por zócalo de GPU o por usuario)
- NVAIE: suscripción anual educativa (por socket GPU o por usuario)
- NVAIE-vGPU: licencia combinada de NVAIE + vGPU

Para obtener más información, visite [la guía de pedidos de NVIDIA](#)



Nota: Para obtener números de pieza y precios específicos, póngase en contacto con el equipo de cuentas de Cisco o con el representante de ventas de NVIDIA.

Paso 3: Opciones de asistencia

Asegúrese de que dispone de:

- Contrato de asistencia de Cisco UCS - Asistencia de Cisco (estándar, mejorada, firma)
- Suscripción a NVIDIA AI Enterprise: incluye compatibilidad con NVIDIA

Configuración de referencia

Ejemplo de orden para la infraestructura de IA:

1. Hardware UCS C885A M8 con GPU NVIDIA H200 (8 x H200)
1. Software: NVIDIA AI Enterprise (8 licencias GPU)
2. Support:
 - Soporte de Cisco: mejorado para el hardware de UCS
 - Suscripción a NVIDIA AI Enterprise (incluye soporte 24/7)
3. Servicios: Servicios profesionales opcionales para la implementación

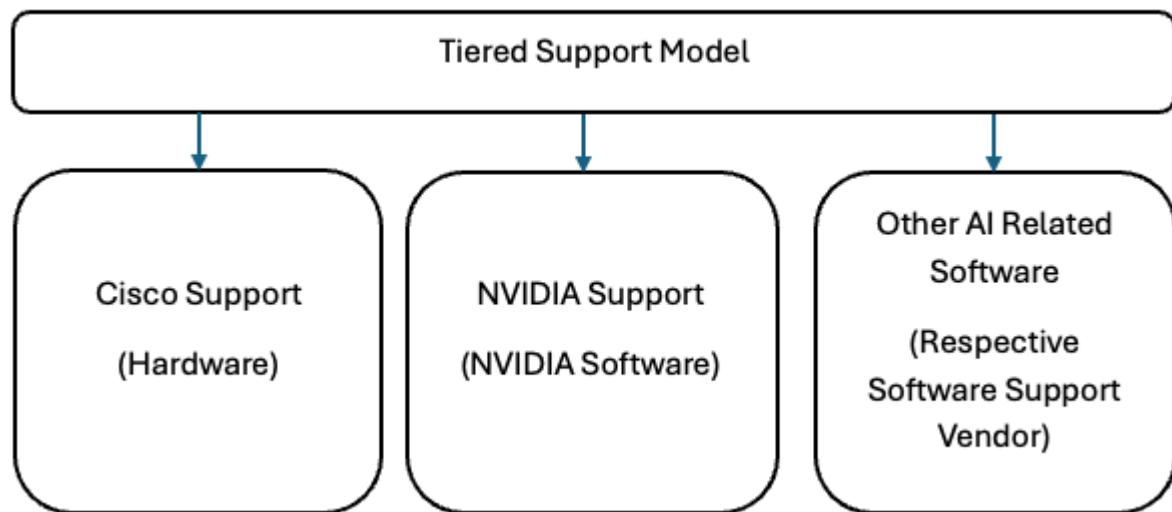
Para obtener guías detalladas sobre pedidos, consulte:

- [Guía de pedidos de NVIDIA AI Enterprise para Cisco UCS](#)

Cómo obtener asistencia

Descripción general del modelo de asistencia

NVIDIA AI Enterprise en Cisco UCS se adhiere a un modelo de soporte por niveles que incluye una clara demarcación entre el soporte de hardware y software:



Nota: Para problemas de software NVIDIA, debe crear un caso de soporte directamente con NVIDIA. Si es necesario, incluya también el servicio de asistencia de Cisco para los esfuerzos conjuntos de resolución de problemas.

Cuándo ponerse en contacto con el servicio de asistencia de Cisco

- 'Problemas del hardware:
 - Errores o fallos del servidor Cisco UCS
 - Diagnóstico y sustitución de hardware
 - Actualizaciones de firmware para componentes de UCS
 - Problemas de configuración de UCS Manager o Intersight
 - Problemas de instalación o detección de GPU física
 - Problemas de conectividad de red con la infraestructura UCS
- Software de Cisco:
 - Problemas de UCS Manager o Cisco Intersight
 - Configuración de red de Cisco para infraestructura de IA

Cómo abrir un caso de soporte de Cisco:

1. Visite: [Contactos de soporte global de Cisco](#)
2. Llame directamente al servicio de asistencia de Cisco (línea directa las 24 horas del día, los 7 días de la semana)
3. Usar el Administrador de casos de soporte de Cisco
4. Contacto mediante Cisco Intersight (para dispositivos conectados a la nube)

Información necesaria antes de crear un caso de soporte:

- Número de serie y modelo de Cisco UCS
- Número de contrato de servicio
- Descripción detallada del problema
- Archivos de soporte técnico (show tech-support)
- Detalles de configuración e inventario de GPU

Cuándo ponerse en contacto con el soporte técnico de NVIDIA

Software NVIDIA AI Enterprise:

- Problemas de instalación o configuración del marco AI
- Preguntas sobre el flujo de trabajo NVIDIA AI
- Implementación de Triton Inference Server
- Problemas con el software NVIDIA vGPU
- Problemas del operador de GPU o del operador de red
- Problemas de tiempo de ejecución del contenedor con NVIDIA Container Toolkit
- Activación o gestión de licencias
- Ajuste del rendimiento para cargas de trabajo de IA
- Errores de software o solicitudes de características
- Preguntas sobre vulnerabilidades de seguridad

Problemas de controladores y software:

- Compatibilidad o instalación del controlador NVIDIA
- problemas del kit de herramientas CUDA
- problemas de cuDNN o TensorRT
- Uso de GPU o problemas de rendimiento

Cómo abrir un caso:

1. Visite el [portal](#) de [asistencia](#) empresarial [de](#) NVIDIA
2. Inicie sesión con su cuenta NVIDIA Enterprise
3. Seleccione "Crear caso" y proporcione:
 - Información de licencia de NVIDIA AI Enterprise
 - Detalles del servidor y la GPU
 - 'Versiones de software'
 - Descripción detallada del problema
 - Registros y resultados de diagnóstico

Acceso al portal de asistencia:

- Requiere una suscripción activa a NVIDIA AI Enterprise
- Configuración de la cuenta durante el aprovisionamiento de licencias
- Acceso a la base de conocimientos, descargas y gestión de casos

Compatibilidad con otro software relacionado con la IA

Cumplir con los procedimientos de soporte proporcionados por el proveedor de software/aplicaciones de IA correspondiente

La información que contiene este documento se creó a partir de los dispositivos en un ambiente de laboratorio específico. Todos los dispositivos que se utilizan en este documento se pusieron en funcionamiento con una configuración verificada (predeterminada). Si tiene una red en vivo, asegúrese de entender el posible impacto de cualquier comando.

Acerca de esta traducción

Cisco ha traducido este documento combinando la traducción automática y los recursos humanos a fin de ofrecer a nuestros usuarios en todo el mundo contenido en su propio idioma.

Tenga en cuenta que incluso la mejor traducción automática podría no ser tan precisa como la proporcionada por un traductor profesional.

Cisco Systems, Inc. no asume ninguna responsabilidad por la precisión de estas traducciones y recomienda remitirse siempre al documento original escrito en inglés (insertar vínculo URL).