

Envolva o suporte para servidores Cisco AI UCS que contêm GPUs Nvidia

Contents

[Introdução](#)

[Informações de Apoio](#)

[Visão geral do software corporativo NVIDIA AI](#)

[O que é NVIDIA AI Enterprise?](#)

[Principais componentes](#)

[Modelos de implantação](#)

[Benefícios de suporte com a licença empresarial NVIDIA AI](#)

[Informações de pedido](#)

[Como obter suporte](#)

[Quando entrar em contato com o suporte da Cisco](#)

[Quando entrar em contato com o suporte NVIDIA](#)

[Suporte para outros softwares relacionados à IA](#)

Introdução

Este documento descreve o NVIDIA AI Enterprise Software (NVAIE) Suite, bem como os benefícios de suporte abrangentes disponíveis quando implementado nas plataformas Cisco Unified Computing System (UCS). Este documento também descreve como utilizar o suporte da Cisco e da NVIDIA para obter a implementação e a assistência operacional ideais.

Informações de Apoio

À medida que a inteligência artificial e as cargas de trabalho de aprendizagem automática se tornam vitais para as operações empresariais, as empresas exigem não apenas uma infraestrutura poderosa, mas também software e suporte de nível empresarial. O software corporativo NVIDIA AI oferece uma plataforma de IA abrangente, segura e compatível, otimizada para implantação na infraestrutura do Cisco UCS.

Este documento descreve:

- O que o software NVIDIA AI Enterprise envolve
- Os benefícios de suporte do licenciamento do pacote de software NVAIE
- Como envolver corretamente os canais de suporte para o NVAIE no Cisco UCS

Visão geral do software corporativo NVIDIA AI

O que é NVIDIA AI Enterprise?

O NVIDIA AI Enterprise (NVAIE) é um pacote de software abrangente que fornece segurança, estabilidade, capacidade de gerenciamento e suporte de nível empresarial para cargas de trabalho de IA executadas em ambientes virtualizados e nativos da nuvem. Ele é otimizado e certificado para ser executado em plataformas corporativas líderes, incluindo o Cisco UCS.

Principais componentes



Note: Para obter uma lista atual e completa, visite [Nvidia AI Enterprise](#)

Estruturas e ferramentas da IA:

- Fluxos de trabalho do NVIDIA AI - Fluxos de trabalho completos e pré-desenvolvidos para aplicações comuns de IA
- Kit de ferramentas NVIDIA TAO - Transfira o kit de ferramentas de aprendizagem para visão computacional, IA conversacional e muito mais
- NVIDIA Riva - SDK para criação de aplicativos de IA por voz
- NVIDIA Morpheus - Estrutura de IA de segurança digital
- NVIDIA Metropolis - plataforma Vision AI para análise de vídeo inteligente
- NVIDIA Clara - Estrutura de aplicativos de IA para serviços de saúde

Software de infraestrutura:

- Software NVIDIA vGPU - Tecnologia de virtualização de GPU (licenciada separadamente ou incluída com NVIDIA)
- Operador de GPU NVIDIA - operador Kubernetes para gerenciamento de GPU
- Operador de rede NVIDIA - Gerencia recursos de rede NVIDIA no Kubernetes
- NVIDIA Container Toolkit - Ferramentas para criar e executar contêineres acelerados por GPU

Pilha de desenvolvimento de IA:

- Estruturas de IA otimizadas (TensorFlow, PyTorch, JAX e assim por diante)
- NVIDIA RAPIDS para ciência de dados
- NVIDIA Triton Inference Server para implantação de IA de produção
- NVIDIA TensorRT para inferência otimizada
- Bibliotecas NVIDIA cuDNN, cuBLAS e CUDA

Gerenciamento e segurança:

- NVIDIA Base Command Manager Essentials - Gerenciamento de infraestrutura para IA
- Sistema de licença NVIDIA - Gerenciamento centralizado de licenças
- Patches de segurança e versões de suporte a longo prazo (LTS)

Modelos de implantação

1. Ambientes virtualizados - VMware vSphere, virtualização Red Hat, GPU virtual NVIDIA
2. Bare Metal - Implantação direta em servidores físicos
3. Nuvem nativa - Kubernetes e implantações em contêineres
4. Híbrida e multi-nuvem - Plataforma de IA consistente entre ambientes

Benefícios de suporte com a licença empresarial NVIDIA AI

Suporte empresarial abrangente

Ao licenciar a NVIDIA AI Enterprise, você recebe benefícios de suporte significativos que vão além do acesso básico ao driver da GPU.

1. Suporte de software de nível de produção
 - Patches e atualizações de software - Atualizações de segurança regulares, correções de bugs e aprimoramentos de recursos
 - Versões de suporte de longo prazo - versões LTS estáveis com janelas de suporte estendidas (até 2 anos)



Note: Os níveis de suporte da Nvidia (básico versus 24 horas por dia, 7 dias por semana) são baseados na licença específica adquirida, conforme descrito [aqui](#)

Componente de serviço	Padrão	24x7
Acesso ao suporte técnico	Horário comercial local	24 horas, 7 dias por semana
Tempos de resposta iniciais de gravidade 1	4 Horário comercial	1 Hora
Tempos de resposta iniciais de gravidade 2	4 Horário comercial	2 horas
Tempos de resposta iniciais de gravidade 3	1 dia útil	4 Horário comercial
Tempos de resposta iniciais de gravidade 4	2 Dias Úteis	1 dia útil
Acesso ao portal do cliente	24x7	24x7
Web	24x7	24x7
Fone	Horário comercial local	24x7

2. Especialização técnica e orientação

- Assistência de implantação de IA - Ajuda com o design de arquitetura e as práticas recomendadas

- Otimização de desempenho - Orientação sobre o ajuste de cargas de trabalho de IA para desempenho máximo
- Suporte de Estrutura - Assistência com TensorFlow, PyTorch e outras estruturas de IA incluídas no conjunto
- Suporte a contêiner e Kubernetes - Ajuda com implantações de IA nativa da nuvem
- Acesso à base de conhecimento - documentação abrangente, guias e recursos de solução de problemas

3. Benefícios de licenciamento e acesso

- Software verificado e validado - versões testadas e certificadas pela empresa
- Portabilidade de licenças - Flexibilidade para mover licenças pela infraestrutura
- Gerenciamento centralizado de licenças - Sistema de licença NVIDIA para gerenciamento de implantações
- Modelo de assinatura - Opções de licenciamento anuais ou plurianuais
- Software Download Access - Portal de download oficial para todos os componentes da NVAIE

4. Segurança e conformidade

- Patches de segurança - Correção oportuna de vulnerabilidades
- Rastreamento CVE - Notificação proativa de problemas de segurança
- Documentação de conformidade - Suporte para requisitos normativos
- Cadeia de fornecimento de software seguro - Componentes de software verificados e assinados

Cobertura de suporte

Recurso	NVIDIA AI Enterprise
Acesso ao driver da GPU	✓
Documentação básica	✓
Fóruns da comunidade	✓
Suite de software AI	✓
Suporte empresarial (24 horas por dia, 7 dias por semana)	✓
Tempos de resposta do SLA	✓

Recurso	NVIDIA AI Enterprise
Patches de segurança	✓
Suporte a longo prazo	✓
Certificação de produção	✓
Gerenciamento técnico de contas	Complemento opcional
Acesso ao fluxo de trabalho de IA	✓
Ferramentas de gerenciamento	✓

Informações de pedido

Cisco UCS com NVIDIA AI Enterprise

Passo 1: Selecione o hardware do Cisco UCS

Escolha a plataforma apropriada do Cisco UCS e a configuração da GPU NVIDIA:

Exemplo de configurações:

- UCS C885A M8 com GPUs NVIDIA HGX H200
- UCS C845A M8 com GPUs NVIDIA RTX6000
- UCS C880A M8 com GPUs NVIDIA HGX B300
- UCS C240 M7 com GPUs NVIDIA L40S
- UCS X-Series com GPU NVIDIA H100 para implantações modulares



Note: Cada NVIDIA H200 NVL, H100 PCIe e H100 NVL GPU vem com uma assinatura NVIDIA Enterprise de cinco anos, que inclui a NVIDIA AI Enterprise. A ativação de software é necessária. Você deve ativar a licença e abrir um caso de suporte de software com a NVIDIA, se necessário.

Clique [aqui](#) para obter mais detalhes.

Por exemplo, UCS C845A M8, UCSX-210C-M8, UCS C240 M7 com GPUs NVIDIA H200 incluem assinatura de cinco anos do NVIDIA Enterprise.

Passo 2: Licença empresarial NVIDIA AI

As licenças NVIDIA AI Enterprise são vendidas separadamente e podem ser adquiridas por meio de:

- Cisco Commerce - junto com as compras do Cisco UCS



Note: A licença NVIDIA também pode ser adquirida fora da Cisco.

Tipos de licença:

- NVAIE- Standard - Assinatura anual padrão (por soquete GPU ou por usuário)
- NVAIE- Educacional - Assinatura anual educacional (por soquete GPU ou por usuário)
- NVAIE-vGPU - licença combinada NVAIE + vGPU

Para obter mais detalhes, visite [o Guia de pedidos da NVIDIA](#)



Note: Para obter preços e números de peças específicos, entre em contato com a equipe de contas da Cisco ou com o representante de vendas da NVIDIA.

Passo 3: Opções de suporte

Verifique se você tem:

- Contrato de suporte do Cisco UCS - Suporte da Cisco (Padrão, Avançado, Assinatura)
- Assinatura empresarial NVIDIA AI - Inclui suporte para NVIDIA

Configuração de referência

Exemplo de pedido de infraestrutura de IA:

1. Hardware: UCS C885A M8 com GPUs NVIDIA H200 (8x H200)
1. Software: NVIDIA AI Enterprise (8 licenças de GPU)
2. Support:
 - Suporte da Cisco - Aprimorado para hardware UCS
 - Assinatura NVIDIA AI Enterprise (inclui suporte 24 horas por dia, 7 dias por semana)
3. Serviços: Serviços profissionais opcionais para implantação

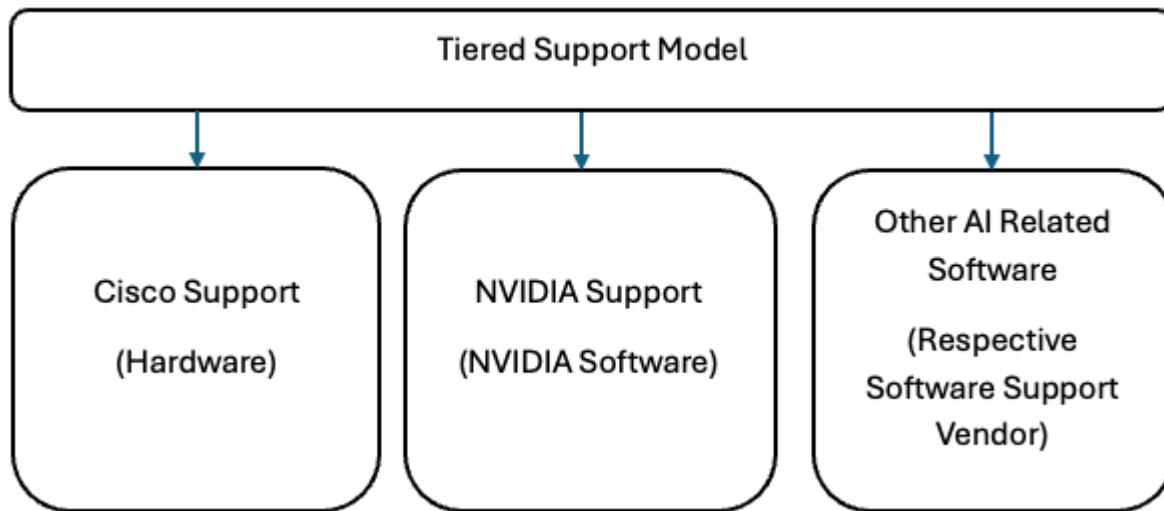
Para obter guias de pedidos detalhados, consulte:

- [Guia de pedidos empresariais NVIDIA AI para Cisco UCS](#)

Como obter suporte

Visão geral do modelo de suporte

A NVIDIA AI Enterprise no Cisco UCS adere a um Modelo de Suporte Hierárquico que inclui uma demarcação clara entre o suporte de hardware e software:



Note: Para problemas de software NVIDIA, você deve criar um caso de suporte diretamente com a NVIDIA. Se necessário, inclua também o Suporte da Cisco para esforços conjuntos de solução de problemas.

Quando entrar em contato com o suporte da Cisco

- Problemas de hardware:
 - Falhas ou erros do servidor Cisco UCS
 - Diagnóstico e substituição de hardware
 - Atualizações de firmware para componentes do UCS
 - Problemas de configuração do UCS Manager ou da Intersight
 - Problemas de instalação ou detecção física da GPU
 - Problemas de conectividade de rede com a infraestrutura do UCS
- Software da Cisco:
 - Problemas do UCS Manager ou do Cisco Intersight
 - Configuração de rede da Cisco para infraestrutura de IA

Como abrir um caso de suporte da Cisco:

1. Visite: [Contatos mundiais de suporte da Cisco](#)
2. Ligue diretamente para o suporte da Cisco (linha direta 24 horas por dia, 7 dias por semana)
3. Use o Cisco Support Case Manager

4. Entre em contato com a Cisco Intersight (para dispositivos conectados à nuvem)

Informações necessárias antes da criação do caso de suporte:

- Modelo e número de série do Cisco UCS
- Número do contrato de serviço
- Descrição detalhada do problema
- Arquivos de suporte técnico (show tech-support)
- Detalhes de configuração e inventário de GPU

Quando entrar em contato com o suporte NVIDIA

Software corporativo NVIDIA AI:

- Problemas de instalação ou configuração da estrutura de IA
- Perguntas sobre o fluxo de trabalho da NVIDIA AI
- Implantação do Triton Inference Server
- Problemas com o software NVIDIA vGPU
- Problemas do operador de GPU ou do operador de rede
- Problemas de tempo de execução de contêiner com o NVIDIA Container Toolkit
- Ativação ou gerenciamento de licenças
- Ajuste de desempenho para cargas de trabalho de IA
- Bugs de software ou solicitações de recursos
- Perguntas sobre vulnerabilidade de segurança

Problemas de driver e software:

- Instalação ou compatibilidade do driver NVIDIA
- Problemas do kit de ferramentas CUDA
- problemas de cuDNN ou TensorRT
- Problemas de utilização ou desempenho da GPU

Como abrir um caso:

1. Visite o [Portal](#) de [Suporte](#) Empresarial NVIDIA
2. Faça login com sua conta NVIDIA Enterprise
3. Selecione "Create Case" (Criar caso) e forneça:
 - Informações de licença NVIDIA AI Enterprise
 - Detalhes do servidor e da GPU
 - Versões de software
 - Descrição detalhada do problema
 - Registros e saída de diagnóstico

Acesso ao portal de suporte:

- Requer assinatura ativa do NVIDIA AI Enterprise
- Configuração de conta durante provisionamento de licença
- Acesso à base de conhecimento, downloads e gerenciamento de casos

Suporte para outros softwares relacionados à IA

Cumprir os procedimentos de suporte fornecidos pelo Fornecedor de Software/Aplicativo de IA aplicável

As informações neste documento foram criadas a partir de dispositivos em um ambiente de laboratório específico. Todos os dispositivos utilizados neste documento foram iniciados com uma configuração (padrão) inicial. Se a rede estiver ativa, certifique-se de que você entenda o impacto potencial de qualquer comando.

Sobre esta tradução

A Cisco traduziu este documento com a ajuda de tecnologias de tradução automática e humana para oferecer conteúdo de suporte aos seus usuários no seu próprio idioma, independentemente da localização.

Observe que mesmo a melhor tradução automática não será tão precisa quanto as realizadas por um tradutor profissional.

A Cisco Systems, Inc. não se responsabiliza pela precisão destas traduções e recomenda que o documento original em inglês ([link fornecido](#)) seja sempre consultado.