

DCAIE

AI Solutions on Cisco Infrastructure Essentials

32 horas

Professional

Cisco

Cisco Continuing Education Credits

34 CE Credits

INTRODUÇÃO

The AI Solutions on Cisco Infrastructure Essentials (DCAIE) training covers the essentials of deploying, migrating, and operating AI solutions on Cisco data center infrastructure.

OBJETIVO DO CURSO

Describe key concepts in AI, machine learning, and deep learning

Describe generative AI, challenges, and future trends

Explain how AI enhances network management and security

Describe key concepts and architecture of AI-ML clusters

Use Jupyter Lab and Generative AI to automate network operations

Describe essential components for AI infrastructure setup

Evaluate workload placement strategies and AI system interoperability

Explore AI compliance standards and governance frameworks

Describe sustainable AI infrastructure practices

Describe key network challenges for AI/ML workloads

Describe role of optical and copper technologies in AI/ML data centers

Describe network connectivity models and designs

Describe Layer 2 and Layer 3 protocols for AI and fog computing

Explain RDMA and RoCE protocols

Understand high-performance Ethernet fabric architecture

Explain QoS tools for lossless RoCE networks

Describe ECN and PFC mechanisms

Describe AI-specific hardware and compute requirements

Use NDFC to configure a fabric optimized for AI/ML workloads

PÚBLICO-ALVO

Network Designers, Network Administrators, Storage Administrators, Network Engineers, Systems Engineers, Data Center Engineers, Technical Solutions Architects, Field Engineers

PRÉ-REQUISITOS

No formal prerequisites. Recommended: familiarity with Cisco UCS, Nexus switch portfolio, and Data Center core technologies.

Course Outline

Fundamentals of AI and Generative AI
AI Use Cases and AI-ML Clusters
AI Toolset (Jupyter Notebook)
AI Infrastructure
AI Workload Placements and Interoperability
AI Policies and Sustainability
AI Infrastructure Design
Key Network Challenges for AI Workloads
AI Transport and Connectivity Models
AI Network Architecture Migration
Application-Level Protocols (RDMA, RoCE)
High Throughput Converged and Lossless Fabrics
Congestive Visibility and Data Performance
AI-Enabling Hardware and Compute Resources
Storage Resources
Setting Up AI Cluster
Deploy and Use Open Source GPT Models for RAG

Lab Outline

AI Toolset - Jupyter Notebook
AI/ML Workload Data Performance
Setting Up AI Cluster
Deploy and Use Open Source GPT Models for RAG