AI Infrastructure – Professional Training

Prepare for NCP-AII Certification

OUTLINE

Training Overview

In today's AI-driven era, the ability to deploy AI clusters efficiently and effectively is crucial for organizations.

This course specializes in optimizing efficiency, reliability, and scalability for deploying AI environments.
It covers various infrastructure aspects, including compute platforms, storage, and the setup and maintenance of AI infrastructure.

The training focuses on key AI building blocks such as GPUs, CPUs, and BlueField networking platform, providing in-depth knowledge and skills to professionals involved in deploying and managing AI infrastructure.

NCP-AII Certification

The **NCP-AI Infrastructure (NCP-AII)** certification is an intermediate-level credential that validates a candidate's ability to deploy, configure, and validate advanced NVIDIA AI infrastructure.

This multi-day workshop is recommended training for the **NCP-AII** certification.

Training Delivery Method

Instructor-led remote training sessions via NVIDIA Teams platform.
Hands-on lab exercises focused on the data center infrastructure.

Target Audience

The course is designed for administrators, DevOps professionals, and IT-related roles who want to gain the knowledge and skills necessary to deploy and maintain AI data centers.

Training Duration

Remote | 5 sessions of 5 hours

Prerequisites

- Knowledge of core networking concepts and principles, including the TCP/IP model, Ethernet standards, basics of routing and switching, common network topologies, and IP addressing schemes.
- Hands-on experience in Linux-like systems administration, such as managing users and permissions, installing software packages, configuring network settings, and troubleshooting common issues in a Linux environment.
- Basic understanding of server hardware components and their roles in a data center environment. This includes knowledge of CPUs, memory, storage devices, and networking interfaces commonly found in servers.
- Knowledge of storage concepts and principles, including different file systems and their characteristics.
- Basic understanding of virtualization technologies, including virtual machines (VMs) and containers. You should be familiar with VM creation, management, and the role of hypervisors in virtualized environments.
- Basic understanding of artificial intelligence (AI) concepts and terminology. This may include knowledge of topics such as machine learning, deep learning, neural networks, and common AI applications.
- Before attending the course, we recommend completing the AI Infrastructure and Operation Fundamentals self-paced course. This course will provide the foundations for this training.

Training Outline

AI in the Data Center Overview

- AI Overview
- Data Center Architecture for AI Workloads

Compute Platforms for AI

- AI Compute Platforms Overview
- Scaling AI Compute
- NVIDIA System Management (NVSM)
- GPU Monitoring with DCGM
- Hardware Installation and Replacement
- NCCL Overview
- AI Cluster Validation
- **Practice 1**: GPU Driver Installation
- **Practice 2:** Data Center Server and GPU Management Tools

Networking for AI

- AI Data Centers Networks

BlueField Networking Platform

- BlueField Overview and Uses Cases
- BlueField Bring-up
- BlueField Interfaces
- **Practice 1:** BlueField Initial Configuration
- **Practice 2:** Inspect The DPU

Storage for AI

- Storage for AI Overview
- Storage Configuration
- **Practice 1**: Testing Storage Performance

AI Data Center Management

- AI Data Center Management Overview
- Base Command Manager (BCM) Overview
    - Head Node Installation
    - Bringing Up the Cluster
- Scheduling AI Workloads with Slurm
- **Practice 1:** Base Command Manager Administration
- **Practice 2**: Slurm Installation

AI Virtualization

- GPU Partitioning
- Configuring MIG - Demo

NVIDIA AI Software

- Using AI Applications with Containers
- **Practice 1:** - Installing and Running AI Software