AI Infrastructure and Operations – Public Training

OUTLINE

Training Overview

In today's AI-driven era, the ability to deploy and manage AI clusters efficiently and effectively is critical for organizations seeking to leverage the power of artificial intelligence. This specialized course is designed to equip professionals with the skills and knowledge necessary to optimize efficiency, reliability, and scalability in deploying AI environments within data centers.

Covering various infrastructure and operation aspects like compute platforms, and storage, the course focuses on key AI components such as GPUs, CPUs, and DPUs.

Participants will gain practical insights into provisioning and managing AI data centers, implementing AI workloads, and utilizing AI virtualization techniques. Through a blend of hands-on training and theoretical learning, participants develop the expertise needed to effectively deploy and manage AI infrastructure, ensuring optimal performance and continuous operation of AI data centers by course end.

Training Delivery Method

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

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 operate AI data centers.

Training Duration

Remote | 8 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
- AI Compute Software
- Operating AI Compute Platforms
- NVIDIA GPU Containers
- Hardware installation – GPUs installation and Validation, Power and Cooling
- Scaling AI Compute
- Validation and Testing – Cables, NCCL, HPL, NeMo
- Monitoring and Troubleshooting NVLink with DCGM
- Troubleshooting Compute Systems Using NVSM

- Practice: Installing the NVIDIA GPU driver and using the nvidia-smi tool
- Practice: Monitoring and Troubleshooting NVLink with DCGM
- Practice: NVIDIA DGX System Health Diagnostics with NVSM

Networking for AI

- AI Data Centers Networks

BlueField Networking Platform for AI

- BlueField Overview and Uses Cases
- BlueField Bring-up + hands-on
    - Installing DOCA
    - Firmware Upgrade
    - Management via RShim
    - BlueField Interfaces – Network Interfaces and OVS Bridges
- BlueField Management
    - Installing DOCA on BlueField
- Running DOCA Services on BlueField
- Practice: Bringing up BlueField
- Practice: Running the BlueMan Service

Storage for AI

- Storage Requirements for AI data Centers
- Storage Architecture for AI
- Storage Configuration, Optimization and Testing
- Practice: Mounting Storage and Testing Performance

AI Data Center Management

- AI Data Center Management Overview
- AI Infrastructure Management and Provisioning with Base Command Manager
    - Overview, Components, Installation, Cluster Management Tools and Troubleshooting
- Scheduling AI Workloads with Slurm

- AI Cluster Orchestration with Kubernetes
- Run:ai Administration
- Practice: Operating and Managing AI Clusters with BCM
- Practice: Scheduling AI Workloads with Slurm
- Practice: Orchestrating Containerized AI Applications with Kubernetes

Virtualizing GPU Resources

- GPU Temporal Partitioning
- GPU Spatial Partitioning
- Virtualizing GPU resources using MIG

NVIDIA AI Software

- Using NGC Containers
- NGC CLI Installation on the Host
- NVIDIA AI Enterprise Software Suite
- Practice: Installing and Running AI Software