Scaling DataOps with NetApp AI Control Plane (NA- DATAOPS)

Course Description

Learn about the features and benefits of NetApp® AI Control Plane by using NetApp DataOps Toolkit and NetApp Astra™ Trident. Discover the advantages of NetApp software solutions to manage AI workloads. Examine machine learning (ML) models by using Jupyter Notebook, Kubeflow pipelines, and KServe. Explore modern transformer models performing generative AI (GenAI) and how to reduce hallucinations by fine-tuning models and implementing a retrievalaugmented generation (RAG) framework. Written with Kubernetes v1.29, Astra Trident 24.06, NetApp DataOps Toolkit 2.5.0, and KubeFlow 1.8.1, this course prepares you to leverage NetApp solutions with artificial intelligence workloads.

Course Duration:

2 days.

Prerequisites:

- Experience with ONTAP, Kubernetes, and Astra Trident.

Objectives:

This course focuses on enabling you to do the following:

- Describe the different software components in the ONTAP architecture
- Describe the advantages of using NetApp AI software solutions to manage inferencing workloads
- Install data operations (DataOps) tools
- Use NetApp DataOps Toolkit to manage storage in inferencing pipelines
- Build generative AI (GenAI) solutions using NetApp solutions
- Discuss how to scale out and scale up inferencing workloads using NetApp solutions

Course Outline

Module 1: Introduction
- Introduction to AI
- Inference modeling
- DataOps
- NetApp and DataOps

Module 2: Installation
- Kubeflow
- NetApp DataOps Toolkit

Module 3: DataOps
- Model lifecycle
- DataOps methodology
- DataOps tasks
- Case study

Module 4: DataOps in generative AI
- Transformer models
- Selecting a transformer model
- Inferencing and modifying output

Module 5: DataOps at scale
- Scaling ML projects
- Hardware solutions
- Cloud solutions and services

Who Should Attend

Systems administrator, architect, and integration engineer