

Updated on 10/02/2026

[Sign Up](#)

SUSE Certified Deployment Specialist Certification Training

3 days (21 hours)

Overview

SUSE AI is a platform for deploying generative artificial intelligence applications on a controlled infrastructure. With RKE2 and Rancher Prime, along with storage, security, and observability components, it provides a foundation for hosting language models and leveraging your data.

Our SUSE Certified Deployment Specialist Certification Training will help you prepare for the official SUSE Certified Deployment Specialist in SUSE AI certification and develop the deployment skills assessed in the exam.

You'll learn how to configure the host system, deploy a Kubernetes cluster, and integrate GPU resources and persistent storage. You'll also install Ollama and Milvus, as well as Open WebUI, to build a cohesive application pipeline.

Upon completion of the training, you will be able to validate a SUSE AI deployment, secure the cluster with SUSE Security, and monitor workloads with SUSE Observability to identify common anomalies.

This preparation combines hands-on workshops and practice questions to reinforce what you've learned, understand the official exam objectives, and organize your review. You'll have the opportunity to take the exam to have your skills certified.

Like all our training courses, this one will introduce you to **the latest stable version** of the technology and its new features.

Objectives

- Identify the components of SUSE AI and the skills assessed by the certification
- Configure the host system and deploy an RKE2 cluster integrated with Rancher Prime
- Install the GPU resources and storage required for AI workloads
- Deploy Ollama and Milvus, as well as Open WebUI
- Install security and observability components to verify the platform's operation
- Prepare for the exam using technical scenarios and practice questions

Target Audience

- Linux System Administrator
- Kubernetes administrator
- DevOps engineer
- AI Infrastructure Engineer
- Infrastructure Consultant

Prerequisites

- Experience installing and administering SUSE Linux Enterprise Server or SUSE Linux Micro
- Experience with Kubernetes for deploying and managing containerized applications
- Proficiency with Linux commands and the fundamentals of networking and storage
- Ability to read YAML manifests and technical documentation in English

Technical Prerequisites

- Computer with 8 GB of RAM recommended for comfortable access to the remote lab
- Up-to-date Firefox or Chrome browser and OpenSSH client available on the computer
- Stable Internet connection allowing access to the lab, registries, and necessary repositories
- Administrator access to a remote lab with virtual machines, persistent storage, and a node equipped with a compatible NVIDIA GPU
- SUSE Linux Enterprise Server or SUSE Linux Micro images compatible with the lab and access accounts for the SUSE components used

SUSE Certified Deployment Specialist Training Agenda

[Day 1 - Morning]

SUSE AI Certification and Architecture

- Overview of the SUSE Certified Deployment Specialist Certification Training and Objectives of the SUSE Certified Deployment Specialist in SUSE AI Certification

- Concepts of generative AI, LLMs, inference, and embeddings
- SUSE AI architecture and the roles of infrastructure components and applications
- Preparing SUSE Linux Enterprise Server or SUSE Linux Micro hosts, networking, and repository access
- Initial sizing of CPU, RAM, GPU, and storage resources based on workloads
- Hands-on workshop: Set up a lab host and create a checklist of deployment prerequisites.

[Day 1 - Afternoon]

Deploying the Kubernetes Cluster

- RKE2 cluster requirements for AI workloads and node role separation
- Installing and configuring server and agent nodes
- Accessing the cluster with kubeconfig and performing essential checks with kubectl
- Importing the cluster into Rancher Prime and verifying its status
- Troubleshooting deployment, connectivity, and service startup issues
- Hands-on Workshop: Deploy an RKE2 cluster, import it into Rancher Prime, and fix a configuration issue.

[Day 2 - Morning]

GPUs, Storage, and Application Deployment

- The role of NVIDIA drivers and the NVIDIA GPU Operator in the cluster
- Installing the GPU Operator and validating exposed GPU resources
- Installing SUSE Storage and preparing persistent volumes
- Configuring StorageClasses and PVCs for AI applications
- Accessing the SUSE Application Collection, authenticating to the registry, and deploying with Helm
- Hands-on Workshop: Enable a GPU node, provision a persistent volume, and verify access to application charts.

[Day 2 - Afternoon]

Language Models and Vector Databases

- Ollama architecture and selecting a model suited to the lab's resources
- Installing Ollama with Helm and configuring model storage
- Loading an LLM, calling inference, and verifying GPU usage
- Principles of RAG, embeddings, and searching in a vector database
- Installing Milvus, dependencies, and availability checks
- Hands-on Workshop: Deploying Ollama and Milvus, running an inference, and testing vector search on a provided dataset.

[Day 3 - Morning]

Conversational Interface and Security

- Installing Open WebUI and configuring its persistence
- Connecting to Ollama, accessing models, and validating exchanges
- Documenting RAG usage and verifying application connections
- Installing SUSE Security in the AI cluster
- Vulnerability checks, network visibility, and workload protection principles
- Hands-on Workshop: Connect Open WebUI to the deployed model, test a document-based interaction, and review SUSE Security controls.

[Day 3 - Afternoon]

Observability and Final Exam Preparation

- Installing the SUSE Observability agent and connecting to the observability instance
- Configuring OpenTelemetry components to monitor GenAI workloads
- Reviewing metrics, traces, and dependencies to analyze an anomaly
- Review of official exam objectives and distinctions between the roles of SUSE AI components
- Response strategy, time management, and reasoned answers to practice questions
- Hands-on workshop: Diagnose an incident on the platform, take a mock exam, and create your personalized study plan.

Target Audience

This training is intended for both individuals and companies—large or small—seeking to train their teams in new, advanced IT technologies or to acquire specific industry knowledge or modern methodologies.

Entry-Level Assessment

The pre-training assessment complies with Qualiopi quality standards. Upon final registration, the learner receives a self-assessment questionnaire that allows us to evaluate their estimated proficiency level across various types of technologies, as well as their expectations and personal goals for the upcoming training, within the limits imposed by the selected format. This questionnaire also allows us to anticipate certain connection or internal security issues within the company (intra-company or virtual classroom) that could pose challenges for monitoring and ensuring the smooth operation of the training session.

Teaching Methods

Hands-On Training: 60% Practical, 40% Theoretical. Training materials distributed in digital format to all participants.

Organization

The course alternates between theoretical input from the instructor—supported by examples and reflection sessions—and group work.

Assessment

At the end of the session, a multiple-choice quiz is administered to verify that participants have successfully acquired the skills.

Certification

A certificate will be issued to each trainee who has completed the entire training program.