EXAM BLUEPRINT GUIDE

Nutanix Certified Professional Artificial Intelligence (NCP-AI) 6.10 Exam

Table of Contents

Author
Jeff Hall, Manager, Technical Certification Development

Contributors
Ashish Sharma, Service Provider Solution Architect
Ashwini Vasanth, Principal Product Manager
Austin Stone, Associate Systems Engineer
Daemon Behr, Advisory Systems Engineer
Drew Plaster, Sr. Network System Administrator
Ganang Saputro, Contingent Worker
Husain Ebrahim, GPT and AI Specialist, EMEA
Ibrahim Ahmad, Solutions Architect
Jai Prakash, Sr. Systems Reliability Engineer
Jaleon Baskerville, Associate Systems Engineer
James Brown, Sr. Staff Consulting Architect
Jesse Gonzalez, Staff Solutions Architect - Cloud Native
John Isdell, Systems Sales Engineer
Josh Stevenson, Sr. Systems Engineer
Kim Mount, Technical Director, Cutter Group
Laura Jordana, Director, Technical Marketing
Logan Grahn, Sr. Associate Systems Engineer
Luke Congdon, Sr. Director, Product Management
Marlon Menezes, Advisory Systems Engineer - Commercial
Maroane Boutayeb, Global Unit Lead - Nutanix on OVHcloud
Nitesh Singh, Sr. Systems Reliability Engineer
Ryan Figler, Sr. Associate Inside Systems Engineer
Venu Vuppalapati, Sr. Staff Global Practice Lead - Database

1. The Exam

1.1 Purpose of Exam

The Nutanix Certified Professional - Artificial Intelligence (NCP-AI) 6.10 exam will measure a candidate's ability to install, configure, optimize, and troubleshoot Nutanix Enterprise AI (NAI), as well as integrate GenAI applications and agents with NAI. Successful candidates demonstrate mastery of these skills and abilities.

1.2 Number of Questions

The NCP-AI 6.10 exam consists of 75 multiple-choice and multiple-response questions.

1.3 Pricing

The cost for the NCP-AI 6.10 exam is $199 USD.

1.4 Passing Score

The passing score for this exam is 3000, using a scaled scoring method. The scale is from 1000-6000. Scaled scores are calculated using a mathematical formula that considers a variety of factors, including the number and type of exam questions included in a specific version of the exam.

Because this combination may vary in different versions of the same examination, scaled scores provide a fair score for everyone based on the version of the exam taken.

1.5 How Objectives Relate to Questions on the Exam

Objectives summarize what the test is designed to measure. Objectives are developed by Exam Developers and Subject Matter Experts based on identified tasks that relate to the job of installing, configuring, optimizing, and troubleshooting Nutanix Enterprise AI (NAI), as well as integrating GenAI applications and agents with NAI.

Once the initial development process is complete, these objectives are verified using an external group of individuals in the actual job role. Finally, a number of questions is determined for each objective, which relates directly to the criticality of the task in the job role.

1.6 Languages

The exam is available in English and Japanese.

1.7 Time Limit

The time limit for the exam is 120 minutes.

1.8 Scheduling and Taking the Exam

The NCP-AI 6.10 exam is delivered via remote proctoring or in-person at select test centers.

If you select remote proctoring, after registering for the exam and providing valid identification, you will receive information on how to take the exam from your location using a web browser. Because the exam is remote proctored, you will be provided with a locked down, monitored, secure exam experience.

If you select in-person testing, you will be able to select a test center near you. On the day of the exam, you will need to arrive at the test center 15 minutes prior to the exam start time with a valid government-issued ID.

1.9 Certification Tracks

The NCP-AI 6.10 exam is a core component of the Nutanix Artificial Intelligence track. Passing this exam results in achieving the NCP-AI 6 certification.

The certification requires a passing score on the exam. While it is not required that you attend a course, Nutanix provides training that covers the objectives on the exam. Details on the recommended training course are provided in Section 4.

1.10 Retake Policy

If a candidate fails an exam on the first attempt, he or she is allowed two additional attempts. There is a seven-day waiting period between attempts. Like the first attempt, these are paid for individually and Nutanix recommends that you allow sufficient time between attempts to be properly prepared and to maximize your chances for success.

Please note: After three attempts, you will be unable to take the exam for 60 days, after which you can email university@nutanix.com and request that your attempts are reset. Nutanix recommends you utilize the time to thoroughly review this guide and the related references and/or take the recommended training for this exam.

1.11 Exam Security

Nutanix reserves the right to refuse certifying a candidate who violates exam security policies. This includes copying and redistribution of exam material, using any type of study material during the exam itself, attempting to photograph exam items and taking an exam using a false identity. Your identity is captured as part of the exam registration process and must be validated before you will be allowed to take the exam.

1.12 Recertification

Once you have passed the Nutanix Certified Professional – Artificial Intelligence 6.10 exam and achieved the NCP-AI 6 certification, it will remain valid for three years.

To maintain your certification status, you must either renew your existing certification, pass an equivalent NCP-level exam within another certification track, or pass the NCM-MCI exam.

1.13 Benefits of Certification

- Digital badge from Credly that you can share on social media
- Access to the Certification store at http://store.nutanix.com for shirts, mugs, and more
- Opportunity to participate as a SME to develop future exams
- Discount on attending Nutanix .NEXT

2. Intended Audience

A candidate for the NCP-AI 6.10 exam and NCP-AI 6 certification has at least three years of virtual infrastructure experience, as well as one year of experience working with cloud native technologies and the Linux command-line interface (CLI).

The successful candidate should have foundational knowledge of virtual infrastructural components, such as virtual machines, hypervisors, and virtual networking. The candidate should also have knowledge of the NCI cloud, cloud-based Infrastructure-as-a-Service (IaaS), GPUs, and Nutanix Unified Storage. Finally, the candidate should possess a Certified Kubernetes Administrator (CKA) certification level of knowledge.

Successful candidates are typically Developers, Support Engineers, Solution Architects, and Cloud Engineers. They may also be AI application owners, Data Scientists, Data Engineers, System Integrators, and/or Value Added Resellers (VARs).

Finally, the successful candidate will most likely have taken training courses, such as the Nutanix Enterprise AI (NAI) Administration (NAIA) course.

3. Objectives Covered in the NCP-AI 6.10 Exam

3.1 Introduction

It is recommended that candidates have the knowledge and skills necessary to install, configure, optimize, and troubleshoot Nutanix Enterprise AI (NAI), as well as integrate GenAI applications and agents with NAI before attempting the NCP-AI 6.10 exam. It is also recommended that the candidate complete the training course described in Section 4 prior to taking the exam.

For the NCP-AI 6 certification, candidates will be tested on the following software versions:

- NAI: version 2.3

3.2 Objectives

Prior to taking this exam, candidates should understand each of the following objectives. Each objective is listed below; along with related tools the candidate should have experience with, and related documentation that contains information relevant to the objective. Please note that some documentation requires access via the Support Portal. Information on creating an account for use with the Support Portal can be found here.

All objectives may also be referenced in other product documentation not specifically highlighted below. The candidate should be familiar with all relevant product documentation or have the equivalent skills.

Section 1 – Deploy a Nutanix Enterprise AI Environment

Objective 1.1: Validate installation prerequisites

Knowledge

- Identify the installation prerequisites
- Identify the installation limitations
- Cite the installation procedure
- Describe the core fundamental components of NAI architecture

References

- Nutanix Enterprise AI Overview
- Nutanix Enterprise AI Requirements
- Nutanix Enterprise AI Limitations
- Nutanix Enterprise AI Deployment Types

- Nutanix Enterprise AI Licensing
- Large Language Models in Nutanix Enterprise AI

Objective 1.2: Install Nutanix Enterprise AI components

Knowledge

- Compare and contrast the installation process for NKP (including app catalog) and non-NKP environments
- Recognize version compatibility between pre-requisite and NAI components
- Perform a dark site installation
- Configure storage classes

References

- Installing or Upgrading Nutanix Enterprise AI on Air-gapped Environments
- Installing or Upgrading Nutanix Enterprise AI
- Downloading Nutanix Enterprise AI
- Nutanix Docker Hub Private Repository
- Installing NKP
- Kubernetes Cluster Setup
- Setting up a Google Kubernetes Engine Cluster
- Creating a Persistent Volume Claim with Nutanix Files
- Creating a Persistent Volume Claim with Nutanix Volumes
- Creating a Storage Class

Objective 1.3: Configure DNS, setup the URL, and manage required certificates

Knowledge

- Identify or implement an FQDN for the NAI installation
- Ensure that the FQDN has a secure certificate
- Validate successful login to UI

References

- Nutanix Enterprise AI Requirements
- Nutanix Enterprise AI Configuration Parameters
- Verifying and Troubleshooting the Domain and Certificate Customization
- Logging into Nutanix Enterprise AI

Section 2 – Configure a Nutanix Enterprise AI Environment

Objective 2.1: Onboard users to Nutanix Enterprise AI

Knowledge

- Differentiate between the user and administrator roles
- Identify the user management operations that can be performed as an administrator
- Given specific scenarios, leverage roles to limit privileges for target users

References

- User Roles in Nutanix Enterprise AI
- Creating Users in Nutanix Enterprise AI
- Activating or Deactivating Users in Nutanix Enterprise AI
- Installing or Upgrading Nutanix Enterprise AI on Air-gapped Environments

Objective 2.2: Import Large Language Models (LLMs)

Knowledge

- Recognize the methods and repos available for importing
- Obtain repo keys for HuggingFace and/or NVIDIA NGC
- Recognize where to add repo keys in the UI for usage
- Explain the manual import process and the requirements

References

- Pre-validated Models
- Importing a Large Language Model from Hugging Face
- Importing a Large Language Model from Hugging Face using Model URL
- Importing a Large Language Model Manually from Hugging Face
- Importing NVIDIA NIMs from NVIDIA NGC Catalog

- Replacing a Hugging Face Token
- Calculating a Large Language Model Size

Objective 2.3: Create endpoints

Knowledge

- Determine a downloaded model to expose via an endpoint
- Determine the number and type of GPUs required for a selected model
- Determine the number of instances required to achieve a certain throughput
- Determine vCPU/memory and inference engine for optimization scenarios

References

- Nutanix Enterprise AI Requirements
- Creating an Endpoint
- Deleting an Endpoint
- Viewing the Endpoint Details
- Endpoint Attributes
- Accessing an Endpoint using Open AI Compatible Clients
- Endpoint Hibernation

Objective 2.4: Create and apply keys for each API endpoint

Knowledge

- Identify the locations to generate and manage API keys
- Identify where to view API keys in an endpoint
- Deactivate an API key
- Add an API key to an existing endpoint

References

- Creating an API Key
- Updating an API Key
- Activating or Deactivating an API Key
- Viewing Sample API Code
- Adding an NVIDIA NGC Personal Key

Objective 2.5: Deliver endpoints to the consumer

Knowledge

- Identify the endpoint URI and model-specific parameters and the API key to be shared with consumers
- Identify tool calling vs non tool calling API curl commands

References

- Viewing Sample Code
- Creating an Endpoint
- Endpoint Attributes
- Widgets in an Endpoint Detail Page
- API Keys in Nutanix Enterprise AI
- Creating an API Key
- Activating or Deactivating an API Key
- Using AnythingLLM with Nutanix Enterprise AI
- Reasoning models on Nutanix Enterprise AI

Section 3 – Perform Day 2 Operations

Objective 3.1: Prepare requirements for connecting the app

Knowledge

- Determine where to get the sample request in the NAI application
- Explain the elements in the sample request and the elements required for the OpenAI-compliant application configuration
- Recognize the different endpoint types and choose the correct one for a given application

References

- Viewing Sample API Code
- Create embeddings
- Viewing Endpoints in Nutanix Enterprise AI
- Accessing an Endpoint using Open AI Compatible Clients

- Creating an API Key
- Testing an Endpoint

Objective 3.2: Interpret performance details and optimize accordingly

Knowledge

- Determine the observability metrics for performance evaluation
- Determine possible changes in resource allocation to remedy latency and throughput issues

References

- Viewing Cluster and Node Usage and Health Statistics
- Cluster Usage and Health Statistics
- Node Usage and Health Statistics

Objective 3.3: Monitor access activity for outlier detection

Knowledge

- Determine where and how to view the top 5 API Keys being used
- Locate the endpoint dashboard and view assigned API keys
- Recognize when to deactivate API keys
- Review and interpret audit events

References

- NAI Dashboard Widgets
- Widgets in an Endpoint Detail Page
- Activating or Deactivating Users in Nutanix Enterprise AI
- Activating or Deactivating an API Key
- Viewing Audit Events in Nutanix Enterprise AI

Objective 3.4: Select the appropriate LLM to optimize output quality

Knowledge

- Determine the prompt input and the LLM output per endpoint to evaluate accuracy through human feedback
- Determine techniques and models that can be used to improve the output quality
- Apply guardrails to improve safety
- Apply rerank models to achieve desired results

References

- Pre-Validated Models
- Imported LLMs Attributes
- Nutanix Enterprise AI Pulse Telemetry

Section 4 – Troubleshoot a Nutanix Enterprise AI Environment

Objective 4.1: Troubleshoot and resolve performance and resource utilization issues

Knowledge

- Determine where to view infrastructure performance
- Recognize how to filter by GPU nodes and review resulting GPU utilization graph to determine which GPUs are being heavily used
- Determine if an endpoint is using GPU
- Recognize which type of GPU an endpoint is using.
- Determine if endpoint is using CPU-based acceleration or not

References

- Cluster Usage and Health Statistics
- Node Usage and Health Statistics
- Viewing Endpoints in Nutanix Enterprise AI
- Viewing the Endpoint Details
- Logging Operator
- Logging Architecture
- Interacting with Nodes and Cluster
- Gather Environment Data for Nutanix NKP Support

- Adding GPU Node Pool to a Nutanix Cluster
- Nutanix Enterprise AI Licensing

Objective 4.2: Remediate health check failures on the cluster

Knowledge

- Debug a cluster health fail visible on NAI UI
- Recognize the different components that can cause health check failures
- Analyze Kubernetes NAI system resources to address health check failures
- Determine which layer of the stack is causing the health check failure
- Based on a health check failure diagnosis, determine an appropriate course of action

References

- Cluster Usage and Health Statistics
- Node Usage and Health Statistics
- Workspace Discovery Commands
- NKP Troubleshooting Guide
- Increasing Memory Limits of the kommander-cm Pod
- Audit Events Summary

Objective 4.3: Troubleshoot model import and endpoint creation

Knowledge

- Identify the failure scenarios where model download fails due to misconfigurations and/or connectivity (e.g., prevalidated, custom, and/or restricted networks)
    - Troubleshoot CSI driver connectivity
    - Ensure model EULA was accepted on HuggingFace (Llama models)
    - Ensure HuggingFace or NVIDIA token is valid
- Determine available allocatable compute resources (e.g., CPU, memory, GPUs, taints) that could prevent endpoints from being scheduled
- Recognize if all prerequisites were successfully installed (e.g. Kserve)
- Diagnose the cause of container images failing to be downloaded or be stored on Kubernetes nodes

References

- Nutanix Enterprise AI Requirements
- Cluster Usage and Health Statistics
- Large Language Models in Nutanix Enterprise AI
- Importing a Large Language Model from Hugging Face
- Importing a Large Language Model Manually from Hugging Face
- Configuring Node Pool Information
- Setting up a Nutanix Kubernetes Platform Cluster
- Creating an Endpoint
- Endpoints Attributes
- Widgets in an Endpoint Details Page
- User Roles in Nutanix Enterprise AI
- Access an Endpoint using Open AI Compatible Clients
- Konvoy Image Builder
- Nutanix Image Builder

Section 5 – Connect Applications to a Nutanix Enterprise AI Environment

Objective 5.1: Configure and validate an application with the endpoint

Knowledge

- Differentiate between model and endpoint types to be consumed by the application
- Recognize the purpose and use case for integrating the various types of models
- Issue a simple query to the OpenAI-compatible NAI API endpoints using Python or Curl
- Investigate and address application integration issues

References

- Pre-Validated Models
- Calculating a Large Language Model Size
- Endpoints in Nutanix Enterprise AI
- Viewing the Endpoint Details
- Accessing an Endpoint using Open AI Compatible Clients
- Testing an Endpoint
- Viewing Audit Events in Nutanix Enterprise AI

Objective 5.2: Check endpoint metrics corresponding to application usage

Knowledge

- Identify the latency and number of API requests per endpoint associated with application
- Describe how to correlate the application with NAI endpoint metrics

References

- Viewing the Endpoint Details
- Widgets in an Endpoint Details Page
- Endpoint Attributes
- Endpoint Hibernation
- Testing an Endpoint
- Viewing API Keys in Nutanix Enterprise AI
- Viewing Audit Events in Nutanix Enterprise AI

4. NCP-AI 6.10 Training Recommendations

4.1 Course Recommendation

Nutanix offers a course that provides training on the objectives tested for in the exam. More information on this course, including delivery methods and pricing, can be found at nutanix.com/training.

The course details are as follows

The **Nutanix® Enterprise AI (NAI) Administration (NAIA)** course equips you with everything you need to confidently roll out NAI in your organization. From selecting the right Kubernetes versions and GPU configurations to setting up storage classes, services, and applications, you'll gain a comprehensive understanding of the deployment process.

Explore how to build and manage Kubernetes clusters across leading platforms like Nutanix Kubernetes Platform (NKP), Amazon EKS, Azure AKS, and Google GKE. Dive into the installation of NAI and master essential administrative tasks such as user management, model imports, endpoint configuration, and API key handling.

You'll also learn how to monitor system performance with real-time metrics, leverage built-in audit tools for transparency, and investigate common issues with confidence.

.

The NAIA course will explore a number of subjects, including:

- Preparing to install NAI
- Installing and logging into NAI
- Post-installation configuration tasks
- Working with models, endpoints, and API keys
- Monitoring performance/health metrics and auditing events

This course is available online or instructor-led. More information including schedules and how to register can be found at www.nutanix.com/university.

The material provided in the course covers a majority of the objectives (approximately 80%) that appear on the NCP-AI 6.10 exam and is recommended for individuals who want to gain a good understanding of these objectives. Please note that additional exposure to a Nutanix environment is highly recommended.

5. Resources

5.1 Nutanix Community Edition

The Nutanix Community Edition is a free product that allows you to deploy a Nutanix Cloud Platform. To download the software and build your own environment for exam preparation, click here.

5.2 Test Drive

You can also take a 2-hour Hyperconverged Test Drive, which utilizes the Nutanix Community Edition, by clicking here.

5.3 The Nutanix Community

Connect with cloud builders from around the world, learn from IT Pros in your industry and share experiences on the Nutanix Community. The community maintains an area focused on Nutanix certifications, which is located here.

5.4 Additional Cloud Native Resources

Find a wealth of additional Nutanix Enterprise AI resources here.