

AI/ML Design Workshop

Product Code: CNS-INF-A-WRK-DES-GPT

At-a-Glance

Stage: Design

The Nutanix Artificial Intelligence/Machine Learning (AI/ML) Design Workshop offers IT teams in-depth and practical guidance to create a comprehensive infrastructure design for AI/ML workloads based on a Generative Pre-trained Transformer (GPT). It covers various aspects such as performance, scalability, flexibility, integration, and operational needs. This workshop is beneficial during the Design stage of the AI/ML solution journey.

Related Services

- AI/ML Planning Workshop
- NAI Deployment
- AI/ML Strategy & Optimization Workshop Series

Service Scope

A series of design workshops is delivered by a highly skilled consultant with strong domain expertise and rich experience to ensure that the solution requirements and required outcomes are identified. The consultant works collaboratively with key customer stakeholders from architecture, virtualization, and networking teams during the design workshop to gather requirements and develop the design. After the Design workshop, the consultant develops an infrastructure Design document for the AI/ML inference platform, a configuration workbook that addresses conceptual, logical, and physical NCI design elements.

This service includes the following activities:

- Gather and document solution requirements, constraints, assumptions, dependencies, and decisions in a series of workshops
- Develop NCI cluster design for AI workloads
- Discuss the GPU selection and configuration options for inference
- Assess the network requirements and design virtual networking, including integration with the physical network
- Validate cluster size and platform selection based on workload details provided by the customer
- Design security including data-at-rest encryption, Secure Sockets Layer (SSL) certificate, password complexity, and syslog

Limitations

- For each quantity purchased, infrastructure design is limited to a single AI/ML inference use case.
- Management and other cluster designs require a separate Infrastructure Design Workshop for each additional cluster

- Excludes Nutanix Kubernetes Platform (NKP), Nutanix Database Service (NDB), and Nutanix Unified Storage (NUS) design
- The design service is limited to a single physical site

Note: For AI/ML workloads running on Bare Metal or Public Cloud, a custom Statement of work is required

Supported Hypervisors

- Nutanix AHV

Prerequisites

- None

Required Product Licenses

- None

Delivered Artifacts

- Configuration Workbook
- Design Document

Duration

Typically up to 5 days

Related Products

- Nutanix Cloud Infrastructure (NCI)
- Nutanix Enterprise AI (NAI)
- Nutanix Kubernetes Platform (NKP)
- Nutanix Unified Storage (NUS)
- Nutanix Database Service (NDB)

Terms and Conditions

This document contains the entire scope of the service offer. Anything not explicitly included above is out of scope. This service offer is subject to the Nutanix Services General Terms and Conditions, which can be viewed at <https://www.nutanix.com/support-services/consulting-services/terms-and-conditions>

©2025 Nutanix, Inc. All rights reserved. Nutanix, the Nutanix logo, and all Nutanix product and service names mentioned herein are registered trademarks or trademarks of Nutanix, Inc. in the United States and other countries. Nutanix, Inc. is not affiliated with VMware by Broadcom or Broadcom. VMware and the various VMware product names recited herein are registered or unregistered trademarks of Broadcom in the United States and/or other countries. All other brand names mentioned herein are for identification purposes only and may be the trademarks of their respective holder(s).