



NVIDIA-Certified Professional: AI Infrastructure 认证考试学习指南



NVIDIA-Certified Professional: AI Infrastructure 认证考试学习指南

目录

系统和网络：	2
考试权重 17%	
故障排查和优化：	3
考试权重 15%	
系统和服务器：	4
考试权重 33%	
物理层管理：	5
考试权重 35%	

针对 NVIDIA-Certified Professional: AI Infrastructure (NCP-AII) 认证考试，本学习指南包含认证所涵盖的各个技术主题的介绍，并推荐了相关培训课程和阅读资料。

查看所有 NVIDIA 认证，[请单击此处](#)。

考试适用人群

AI 基础设施专业人员，负责完成系统、网络和存储领域的一系列任务。主要业务领域包括配置网络拓扑，安装基于 GPU 的服务器和优化存储系统。AI 基础设施专业人员需能够设置 NVIDIA® NVLink™ 交换机，配置 InfiniBand 硬件，并安装 NVIDIA Bluefield® DPU 服务器。包括排除硬件故障，更换故障组件，以及优化 AMD 和英特尔服务器的性能。物理层管理包括安装 NVIDIA GPU 驱动，为 AI 工作负载配置 NVIDIA 多实例 GPU (MIG)，以及部署 DPU 操作系统映像。持续优化，确保 AI 基础设施长期高效运行并能适应不断变化的工作负载需求。

相关工作职责示例

1. 为 AI 和高性能计算 (HPC) 工作负载安装并配置 NVIDIA GPU
2. 在数据中心部署和管理用于 AI 研究和应用的 NVIDIA DGX™ 系统
3. 实施 NVIDIA AI Enterprise 软件套件，实现数据中心现代化并加速 AI 应用
4. 配置和优化 NVIDIA CUDA® 库，加速深度神经网络
5. 安装和维护用于容器化 AI 工作负载的 NVIDIA GPU 驱动和 NVIDIA Container Toolkit
6. 设置和管理 NVIDIA BlueField DPU，增强 AI 环境中的网络、存储和安全
7. 为 AI 和 HPC 工作负载实施和配置 NVIDIA MIG 技术
8. 利用 NVIDIA NGC™ 来访问和部署优化的 AI 软件和容器
9. 将 NVIDIA AI 基础设施与 Docker 和 Kubernetes 等云原生技术集成
10. 优化存储解决方案和网络，支持 NVIDIA GPU 加速 AI 工作流
11. 实施和管理 NVIDIA AI Enterprise 软件，实现端到端 AI 开发和部署

建议具备的知识和经验

1. NVIDIA GPU/DPU 技术、AI 软件堆栈和数据中心管理专业知识，支持高性能 AI 工作负载
2. 理工科背景，如计算机科学、软件工程、AI 等

认证主题和参考资料

系统和网络：考试权重 17%

设计和实施数据中心网络基础设施，包括确定合适的网络拓扑、配置网络端口以确保主机连接、安装和设置用于高速通信的 NVLink Switch，以及在 InfiniBand 和 NVIDIA 网络硬件上配置路由表以优化数据流和连接。

- 1.1 确定数据中心的网络拓扑
- 1.2 为主机设置网络端口
- 1.3 安装和配置 NVIDIA NVLink Switch
- 1.4 配置 InfiniBand 和 NVIDIA 网络硬件上的路由表

培训推荐（可选）

课程：AI Infrastructure Professional Public Training ([了解详情](#))，模组 3，AI 网络

- > 第 1 单元：AI 数据中心网络
- > 第 2 单元：构建和维护 InfiniBand AI 基础设施
- > 第 4 单元：AI 数据中心管理

阅读内容推荐

- > [NVIDIA System Management Interface](#)
- > [NVIDIA System Management User Guide](#)
- > [NVIDIA CUDA Compiler Driver NVCC](#)
- > [NVIDIA Virtual GPU Software](#)
- > [InfiniBand Fabric Utilities](#)
- > [Getting Started With the NGC Command-Line Interface \(CLI\)](#)
- > [NVIDIA LinkX® Cables and Transceivers](#)

故障排查和优化：考试权重 15%

判断和解决硬件问题 (例如 GPU、风扇、网卡、电源)，优化 AMD 和英特尔服务器的存储和性能，以及更换故障组件以确保系统效率和可靠性。

2.1 优化存储

2.2 识别和排除硬件故障，包括 CPU、风扇和网卡

2.3 识别故障网卡、GPU 和电源

2.4 更换故障网卡、GPU 和电源

培训推荐（可选）

课程：AI Infrastructure Professional Public Training ([了解详情](#))，模组 4，AI 存储

> 第 1 单元：AI 存储概述

> 第 2 单元：AI 存储架构

课程：AI Infrastructure Professional Public Training ([了解详情](#))，模组 2，计算平台概述

> 第 1 单元：AI 计算平台概述

阅读内容推荐

> [NVIDIA System Management User Guide](#)

> [RAPIDS™ cuDF Accelerates pandas Nearly 150x With Zero Code Changes | NVIDIA Technical Blog](#)

> [Choosing the Right Storage for Enterprise AI Workloads](#)

> [NVIDIA ConnectX® Card Replacement](#)

> [NVIDIA DCGM](#)

> [Data Center GPU Manager Guide](#)

> [Introduction to NVIDIA DGX H100/H200 Systems](#)

> [Spotlight: NVIDIA BlueField DPUs Power the VAST Data Platform for AI Workload Optimization](#)

> [NVIDIA-Certified Systems™](#)

系统和服务器：考试权重 33%

检查电源和冷却系统，安装和验证基于 GPU 和 BlueField DPU 的服务器，正确安装硬件，以及识别线缆类型和收发器并保障连接。

3.1 检查电源和冷却系统

3.2 安装基于 GPU 的服务器

3.3 验证已安装的硬件

3.4 安装基于 Bluefield DPU 的服务器

3.5 识别线缆类型和收发器

培训推荐（可选）

课程：AI Infrastructure Professional Public Training ([了解详情](#))，模组 6，AI 数据中心管理

> 第 2 单元：使用 NVIDIA Base Command™ Manager 进行基础设施调配和管理

课程：AI Infrastructure Professional Public Training ([了解详情](#))，模组 5，BlueField DPU

> 第 1 单元：BlueField DPU 概述和用例

> 第 2 单元：BlueField DPU 启动

> 第 3 单元：BlueField DPU 接口

阅读内容推荐

> [NVIDIA LinkX Cables and Transceivers](#)

> [NVIDIA CUDA Compiler Driver NVCC](#)

> [NVIDIA CUDA Compiler Driver NVCC](#)

> [Using Custom DGX Software Utilities for the DGX Station](#)

> [NVIDIA DGX H100/H200 User Guide](#)

> [NVIDIA BlueField-2 Ethernet DPU User Guide](#)

> [How Can You Protect Yourself When Working With Computer Components?](#)

> [AI Supercluster: NVIDIA DGX H100 Power and Cooling](#)

> [NVIDIA DGX A100 System](#)

> [Using the NVSM CLI—NVIDIA System Management User Guide 24.03 Documentation](#)

> [Dial It In: Data Centers Need New Metric for Energy Efficiency | NVIDIA Blog](#)

> [System Health Checks and Debugging—NVIDIA DGX SuperPOD™](#)

> [Cooling and Airflow Optimization—NVIDIA DGX SuperPOD](#)

> [NVIDIA BlueField-3 SNAP for NVMe and Virtio-blk v4.2.1](#)

> [NVIDIA DGX OS 5 User Guide](#)

> [NVIDIA DGX SuperPOD with NetApp Design Guide](#)

> [Protecting Sensitive Data and AI Models With Confidential Computing | NVIDIA Technical Blog](#)

物理层管理：考试权重 35%

故障排查是保持最佳性能的关键。职责包括解决 Docker、NVlink/NVswitch 系统的 Fabric Manager 服务、Base Command Manager、Magnum IO 组件及存储性能相关问题。确保这些要素无缝运行对于高效运行 AI 工作负载并防止潜在中断至关重要。

- 4.1 安装、更新和删除 NVIDIA GPU 驱动
- 4.2 安装 NVIDIA Container Toolkit
- 4.3 演示如何将 NVIDIA GPU 与 Docker 配合使用
- 4.4 在主机上安装 NGC CLI
- 4.5 配置和管理 Bluefield 网络平台
- 4.6 为 AI 和 HPC 配置 MIG
- 4.7 将 DPU 操作系统映像部署到 Arm

培训推荐（可选）

课程：AI Infrastructure Professional Public Training ([了解详情](#))，模组 2，AI 计算平台

- > 第 1 单元：AI 计算平台概述
- > 第 3 单元：使用 NVSM 管理 AI 计算平台

课程：AI Infrastructure Professional Public Training ([了解详情](#))，模组 4，AI 存储

- > 模组 5：BlueField DPU

课程：AI Infrastructure Professional Public Training ([了解详情](#))，模组 7，GPU 资源虚拟化

- > 第 1 单元：GPU 资源虚拟化

课程：AI Infrastructure Professional Public Training ([了解详情](#))，模组 1，数据中心 AI 简介

- > 第 2 单元：面向 AI 工作负载的数据中心架构

阅读内容推荐

- > [Getting Started With the NGC CLI](#)
- > [User Guide: NVIDIA AI Enterprise Documentation](#)
- > [Advanced GPU Configuration \(Optional\)—NVIDIA AI Enterprise: VMware Deployment Guide](#)
- > [Installing the NVIDIA Container Toolkit](#)
- > [NVIDIA Data Center Drivers: NVIDIA Data Center GPU Driver Documentation](#)
- > [NVIDIA Multi-Instance GPU User Guide r560](#)
- > [Containers for Deep Learning Frameworks User Guide—NVIDIA Docs](#)
- > [NVIDIA DOCA™ Installation Guide for Linux](#)
- > [NVIDIA DOCA Virtual Functions User Guide](#)
- > [BlueField OOB Ethernet Interface—NVIDIA Docs](#)
- > [NVIDIA BlueField DPU BSP](#)

- > [Host-Side Interface Configuration—NVIDIA Docs](#)
- > [MIG User Guide](#)
- > [Useful nvidia-smi Queries](#)
- > [NGC CLI Installation](#)

遇到问题？

问题咨询，请发邮件至 dlichina@nvidia.com

© 2025 NVIDIA CORPORATION 及其关联公司。保留所有权利。NVIDIA、NVIDIA 徽标、BASE COMMAND、BLUEFIELD、CONNECTX、CUDA、DGX、DGX SUPERPOD、DOCA、LINKX、NGC、NVIDIA 认证系统、NVLINK 和 RAPIDS 均为 NVIDIA CORPORATION 及其关联公司在美国和其他国家 / 地区的商标和 / 或注册商标。其他公司名称和名称可能是各个相应所有者的商标。3929950。2025 年 3 月

