



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



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

目录

考试适用人群	1
相关工作职责示例	1
建议具备的知识和经验	1
安装和部署:	2
考试权重 31%	
管理:	4
考试权重 23%	
工作负载管理:	5
考试权重 23%	
故障排查和优化:	6
考试权重 23%	

针对 NVIDIA-Certified Professional: AI Operations 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 运营专业人士，负责管理和维护数据中心内的 AI 基础架构。监管 AI 计算平台，管理软件和容器，为 AI 工作负载配置网络，确保存储和资源高效利用。需要具备 AI 硬件和软件以及数据中心技术方面的专业知识，以优化性能并支持 AI 驱动应用。参与过集群管理、工作负载调度，并实施虚拟化技术，更大限度地提高 AI 环境的效率。

相关工作职责示例

1. 管理和维护 AI 计算平台，包括 GPU 和其他专用硬件
2. 安装和配置 GPU 驱动和软件
3. 监管 AI 软件堆栈和工具，包括深度学习框架和数据科学库
4. 实施和管理容器，如 Docker 和 NVIDIA NGC™
5. 为 AI 工作负载配置和优化网络基础设施，包括 InfiniBand 和以太网
6. 管理 AI 数据存储解决方案，考虑性能和容量需求
7. 部署和管理数据处理器 (DPU)，加速数据中心工作负载
8. 监控和管理 AI 集群运行状况和资源利用率
9. 实施工作负载管理和调度工具，例如 Slurm 和 Kubernetes
10. 确保 AI 基础架构的高效供电和冷却，确保最佳运行条件

建议具备的知识和经验

1. NVIDIA GPU 和 DPU 技术、AI 软件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

安装和部署：考试权重 31%

评估使用 NVIDIA Mission Control 工具套件，特别是 Base Command Manager (BCM) 的熟练程度。监控、管理和维护集群的实践技能，包括作业调度、资源分配、安全和故障排查。具备部署和配置 Kubernetes、SLURM 和 DOCA 等基本服务的能力。

- 1.1 描述 Mission Control 工具套件
- 1.2 使用 BCM 的 Base View 界面实时监控集群性能、资源利用率和节点运行状况。
- 1.3 使用 BCM 的工作负载管理器 (例如，SLURM 或 Kubernetes) 管理作业调度和资源分配
- 1.4 使用 BCM 在集群节点之间应用补丁、更新固件和同步软件映像
- 1.5 管理用户帐户、角色和权限，确保使用 BCM 安全访问集群
- 1.6 使用 BCM 配置和监控集群节点、DPU 和交换机的网络设置
- 1.7 使用 BCM 诊断和解决集群问题，例如作业失败、节点中断或资源瓶颈
- 1.8 使用 BCM 根据硬件或工作负载要求将计算节点组织和配置为多个类别
- 1.9 使用 BCM 维护文档并生成有关集群使用情况、性能和问题的报告
- 1.10 使用 BCM 在 NVIDIA 主机上安装 Kubernetes 并进行初始化
- 1.11 在 DPU ARM 上部署 DOCA 服务
- 1.12 安装 Run:ai
- 1.13 安装 Slurm

培训推荐 (可选)

课程推荐：[AI Operations \(了解详情\)](#)

- > 模组 2，计算，第 5 单元：NVIDIA GPU 容器
- > 模组 4，AI 存储，第 1 单元：AI 存储概述
- > 模组 5，NVIDIA BlueField® DPU，第 3 单元：在 BlueField DPU 上运行 DOCA 服务
- > 模组 6，AI 数据中心管理，第 2 单元：使用 Base Command Manager 进行基础设施调配和管理

阅读内容推荐

- > [Base Command Manager Documentation](#)
- > [DCGM Diagnostics](#)
- > [Deploy Slurm](#)
- > [Slurm Documentation](#)
- > [NVIDIA Base Command™ Manager](#)
- > [Troubleshooting—NVIDIA Container Toolkit](#)
- > [NVIDIA Multi-Instance GPU User Guide](#)
- > [Base Command Manager Cloud Bursting Manual](#)
- > [NVIDIA BlueField DPU Container Deployment Guide](#)
- > [Deploy NGC Containers on Your Cloud Virtual Machine](#)

- > [About the NVIDIA GPU Operator](#)
- > [How to Build Your GPU Cluster: Process and Hardware Options](#)
- > [Run:ai Docs](#)
- > [Workload Isolation](#)
- > [NVIDIA DGX SuperPOD: Deployment Guide](#)
- > [NVIDIA System Management User Guide](#)

管理：考试权重 23%

管理各种 AI 工作负载环境的熟练程度。管理 Slurm 和 Kubernetes 集群等任务，对于调度和编排任务至关重要。具备管理 AI 工作负载管理平台 Run:ai 以及配置 MIG (多实例 GPU) 的知识。具备处理 AI 工作负载的数据中心架构的能力。

2.1 管理 Slurm 集群

2.2 描述面向 AI 工作负载的数据中心架构

2.3 管理 Run:ai

2.4 管理 Kubernetes

2.5 为 AI 和 HPC 配置 MIG

培训推荐（可选）

课程推荐：[AI Operations \(了解详情\)](#)

- > 模组 6，AI 数据中心管理，第 2 单元：面向 AI 工作负载的数据中心架构
- > 模组 6，AI 数据中心管理，第 3 单元：使用 Slurm 调度 AI 工作负载
- > 模组 6，AI 数据中心管理，第 5 单元：借助 Fleet Command 管理边缘 AI
- > 模组 7，GPU 资源虚拟化

阅读内容推荐

- > [Advanced Cluster Configuration](#)
- > [NVIDIA DGX Cloud Create Documentation](#)
- > [NVIDIA Multi-Instance GPU User Guide](#)
- > [Monitoring GPUs in Kubernetes with DCGM | NVIDIA Technical Blog](#)
- > [NVIDIA Spectrum-X™ Ethernet networking platform](#)
- > [NVIDIA Cloud Native Technologies](#)
- > [Kubernetes Documentation: Administer a Cluster](#)
- > [Debugging and Troubleshooting — NVIDIA DCGM Documentation latest documentation](#)
- > [DCGM Diagnostics](#)
- > [NVIDIA System Management Interface](#)

工作负载管理：考试权重 23%

软硬件安装和配置。包括安装和配置 BCM，并使用 BCM 在 NVIDIA 主机上安装 Kubernetes。需能够部署来自 NGC 的容器以及云虚拟机映像 (VMI) 容器。此外，了解 AI 数据中心的存储需求并在 DPU-Arm 上部署 NVIDIA DOCA™ 服务。

3.1 使用 Kubernetes 部署推理工作负载

3.2 使用 Run:ai 部署推理工作负载

3.3 使用 Slurm 部署训练工作负载

3.4 使用 Run:ai 部署训练工作负载

3.5 使用系统管理工具排查问题

3.6 使用 Run:ai、Slurm 和 Kubernetes 在团队之间分配资源

3.7 部署来自 NGC 的容器

培训推荐（可选）

课程推荐：AI Operations ([了解详情](#))

- > 模组 2，计算，第 1 单元：AI 计算平台概述
- > 模组 2，计算，第 3 单元：使用 NVSM 管理计算系统
- > 模组 2，计算，第 4 单元：使用 DCGM 运行 AI 计算平台
- > 模组 6，AI 数据中心管理，第 2 单元：使用 Base Command Manager 进行基础设施调配和管理
- > 模组 6，AI 数据中心管理，第 4 单元：使用 Kubernetes 进行 AI 集群编排

阅读内容推荐

- > [DOCA Blueman Service Overview](#)
- > [Deploying BlueField Software Using BFB with PXE](#)
- > [DOCA Telemetry Service](#)
- > [Base Command Platform](#)
- > [Base Command Manager Installation Manual](#)
- > [Cloud Overview — NVIDIA AI Enterprise: Cloud Deployment Guide](#)
- > [NGC Catalog User Guide - NVIDIA Docs](#)
- > [NGC Private Registry User Guide - NVIDIA Docs](#)
- > [Deploy Kubernetes — NVIDIA DGX BasePOD](#)
- > [HPC SDK Container Guide 25.7 documentation](#)
- > [Docker Swarm vs Kubernetes: Which is Better in 2025?](#)
- > [NGC Catalog User Guide - NVIDIA Docs](#)
- > [Deploy Kubernetes — NVIDIA DGX BasePOD](#)
- > [Deploy NGC Containers on Your Cloud Virtual Machine - NVIDIA Docs](#)
- > [Accelerating AI Storage by up to 48% with NVIDIA Spectrum-X Networking Platform and Partners](#)
- > [Choose the Right Storage for Enterprise AI Workloads | NVIDIA Technical Blog](#)
- > [Deploy Kubernetes — NVIDIA DGX BasePOD](#)
- > [8 Key Data Storage Requirements for AI You Need to Know](#)

故障排查和优化：考试权重 23%

故障排查是保持最佳性能的关键。具备解决 Docker、NVIDIA NVLink™ 和 NVSwitch™ 系统的 Fabric Manager 服务、BCM、Magnum IO™ 组件及存储性能相关问题的能力。确保组件无缝协同工作，对于 AI 工作负载的高效且稳定运行至关重要。

4.1 Docker 故障排查

4.2 对 NVLink 和 NVSwitch 系统的 Fabric Manager 服务进行故障排查

4.3 Base Command Manager 故障排查

4.4 Magnum IO 组件故障排查

4.5 排查存储性能问题

4.6 对来自 NGC 的容器部署进行故障排查

培训推荐（可选）

课程推荐：[AI Operations \(了解详情\)](#)

- > 模组 2，计算，第 5 单元：NVIDIA GPU 容器
- > 模组 3，AI 网络，第 2 单元：运行 InfiniBand Fabrics
- > 模组 3，AI 网络，第 4 单元：使用 NVIDIA United Fabric Manager (UFM®) 监控 AI 数据中心
- > 模组 4，AI 存储，第 1 单元：AI 存储概述
- > 模组 6，AI 数据中心管理，第 2 单元：使用 Base Command Manager 进行基础设施调配和管理

阅读内容推荐

- > [NVIDIA Fabric Manager User Guide](#)
- > [DCGM User Guide](#)
- > [Slurm Documentation](#)
- > [NV Debug Guide](#)
- > [NVIDIA NGC User Guide](#)
- > [NGC Deploy Containers](#)
- > [Deep Learning Frameworks User Guide](#)
- > [NVIDIA Container Toolkit Troubleshooting Guide](#)
- > [Top Five Most Common Issues with Docker and How to Solve Them](#)
- > [NVIDIA GPU Direct Storage Best Practices Guide](#)
- > [GPU Direct Storage Troubleshooting Guide](#)
- > [Optimizing Data Movement in GPU Apps with Magnum IO Developer Environment](#)
- > [NVIDIA Base Command Manager Documentation](#)

更多咨询

请邮件至 dlichina@nvidia.com