Certs4Sale
Pass With 100% Guarantee

NVIDIA NCP-AIO

NVIDIA AI Operations

Thank You for Downloading NCP-AIO Updated Exam Questions

https://certs4sale.com/NVIDIA/ncp-aio-pdf-exam-dumps

Version: 4.1

Question: 1

A Slurm user needs to submit a batch job script for execution tomorrow.

Which command should be used to complete this task?

A. sbatch -begin=tomorrow

B. submit -begin=tomorrow

C. salloc -begin=tomorrow

D. srun -begin=tomorrow

Answer: A

Explanation:

Comprehensive and Detailed Explanation From Exact Extract:

In Slurm cluster administration, the command to submit a batch job script is sbatch. This command schedules the job to be executed by the Slurm workload manager. The option -begin=tomorrow (or --begin=tomorrow) specifies the start time for the job execution, which in this case is set for tomorrow. The other commands have different purposes:

submit is not a valid Slurm command.

salloc is used to allocate resources interactively but does not submit batch jobs for scheduled execution.

srun runs jobs immediately on allocated resources but is typically used to launch tasks in an active job or interactively, not for batch job submission.

Therefore, the correct command to submit a batch job script for future execution is sbatch -begin=tomorrow.

Question: 2

You are configuring networking for a new AI cluster in your data center. The cluster will handle large-scale distributed training jobs that require fast communication between servers.

What type of networking architecture can maximize performance for these AI workloads?

A. Implement a leaf-spine network topology using standard Ethernet switches to ensure scalability as more nodes are added.

B. Prioritize out-of-band management networks over compute networks to ensure efficient job scheduling across nodes.

C. Use standard Ethernet networking with a focus on increasing bandwidth through multiple connections per server.

D. Use InfiniBand networking to provide low-latency, high-throughput communication between servers in the cluster.

Answer: D

Explanation:

Comprehensive and Detailed Explanation From Exact Extract:

For large-scale AI workloads such as distributed training of large language models, the networking infrastructure must deliver extremely low latency and very high throughput to keep GPUs and compute nodes efficiently synchronized. NVIDIA highlights that InfiniBand networking is essential in AI data centers because it provides ultra-low latency, high bandwidth, adaptive routing, congestion control, and noise isolation—features critical for high-performance AI training clusters.

InfiniBand acts not just as a network but as a computing fabric, integrating compute and communication tightly. Microsoft Azure, a leading cloud provider, uses thousands of miles of InfiniBand cabling to meet the demands of their AI workloads, demonstrating its importance. While Ethernet-based solutions like NVIDIA's Spectrum-X are emerging and optimized for AI, InfiniBand remains the premier choice for AI supercomputing networks.

Therefore, for maximizing performance in a new AI cluster focused on distributed training, InfiniBand networking (option D) is the recommended architecture. Other Ethernet-based approaches provide scalability and bandwidth but cannot match InfiniBand's specialized low-latency and high-throughput performance for AI.

Question: 3

A system administrator needs to optimize the delivery of their AI applications to the edge.

What NVIDIA platform should be used?

A. Base Command Platform

B. Base Command Manager

C. Fleet Command

D. NetQ

Answer: C

Explanation:

Comprehensive and Detailed Explanation From Exact Extract:

NVIDIA Fleet Command is the platform designed specifically to optimize and manage the deployment and delivery of AI applications at the edge. It enables secure and scalable orchestration of AI workloads across distributed edge devices, providing lifecycle management, remote monitoring, and updates. Fleet Command facilitates running AI applications closer to where data is generated (edge), improving latency and operational efficiency.

Base Command Platform and Base Command Manager primarily target data center and AI cluster management for configuration, monitoring, and troubleshooting.

NetQ is focused on network telemetry and network state monitoring rather than application delivery.

Therefore, for AI application delivery and optimization at the edge, Fleet Command is the recommended NVIDIA platform.

Question: 4

A Slurm user is experiencing a frequent issue where a Slurm job is getting stuck in the “PENDING” state and unable to progress to the “RUNNING” state.

Which Slurm command can help the user identify the reason for the job’s pending status?

A. sinfo -R

B. scontrol show job <jobid>

C. sacct -j <job[.step]>

D. squeue -u <user_list>

Answer: B

Explanation:

Comprehensive and Detailed Explanation From Exact Extract:

The Slurm command scontrol show job <jobid> provides detailed information about a specific job, including its current status and, crucially, the reason why a job might be pending. This command shows job details such as resource requirements, dependencies, and any issues blocking the job from running.

sinfo -R displays information about nodes and their reasons for being in various states but does not provide job-specific reasons.

sacct -j shows accounting data for jobs but typically does not explain pending causes.

squeue -u lists jobs by user but does not detail the pending reasons.

Hence, scontrol show job <jobid> is the appropriate command to diagnose why a Slurm job remains in the pending state.

Question: 5

You are a Solutions Architect designing a data center infrastructure for a cloud-based AI application that requires high-performance networking, storage, and security. You need to choose a software framework to program the NVIDIA BlueField DPUs that will be used in the infrastructure. The framework must support the development of custom applications and services, as well as enable tailored solutions for specific workloads. Additionally, the framework should allow for the integration of storage services such as NVMe over Fabrics (NVMe-oF) and elastic block storage.

Which framework should you choose?

A. NVIDIA TensorRT

B. NVIDIA CUDA

C. NVIDIA NSight

D. NVIDIA DOCA

Answer: D

Explanation:

Comprehensive and Detailed Explanation From Exact Extract:

NVIDIA DOCA (Data Center Infrastructure-on-a-Chip Architecture) is the software framework designed to program NVIDIA BlueField DPUs (Data Processing Units). DOCA provides libraries, APIs, and tools to develop custom applications, enabling users to offload, accelerate, and secure data center infrastructure functions on BlueField DPUs.

DOCA supports integration with key data center services including storage protocols such as NVMe over Fabrics (NVMe-oF), elastic block storage, and network security and telemetry. It enables tailored solutions optimized for specific workloads and high-performance infrastructure demands.

TensorRT is focused on AI inference optimization.

CUDA is NVIDIA’s GPU programming model for general-purpose GPU computing, not for DPUs.

NSight is a development environment for debugging and profiling NVIDIA GPUs.

Therefore, NVIDIA DOCA is the correct framework for programming BlueField DPUs in a data center environment requiring custom application development and advanced storage/networking

integration.

THANK YOU FOR DOWNLOADING NCP-AIO UPDATED EXAM QUESTIONS

Note: Thanks For Trying The Demo Of Our NCP-AIO Exam Product

Visit Our Site to Purchase the Full Set of Actual NCP-AIO Exam Questions With Answers.

Money Back Guarantee

Click The Link Below

https://certs4sale.com/NVIDIA/ncp-aio-pdf-exam-dumps