Download NVIDIA NCP-AIN Exam Dumps Free

Shared by West on 17-06-2026

For More Free Questions and Preparation Resources

Check the Links on Last Page

Question 1

Question Type: MultipleChoice

[Spectrum-X Optimization]

Your organization is planning to utilize Ethernet for an upcoming AI project. Spectrum-X is the selected platform for this deployment, and Adaptive Routing is a key feature.

What are the requirements included in the Spectrum-X RA for adaptive routing?

Options:

A- SN4700, BlueField-3 SuperNIC, DDR, RoCE traffic
B- SN5600, BlueField-3 SuperNIC, DDR, RoCE traffic
C- SN5600, BlueField-3 SuperNIC, DDR, TCP traffic

Answer:

B

Explanation:

The NVIDIA Spectrum-X Reference Architecture (RA) 1.0.1 is designed for Ethernet AI cloud deployments and includes the SN5600 Spectrum-4 switches and BlueField-3 SuperNICs. This architecture supports adaptive routing and DOCA programmable congestion control (PCC) for lossless RoCE traffic, optimizing performance for AI workloads.

The SN5600 switch offers 64 ports of 800GbE in a dense 2U form factor, providing high throughput and low latency essential for AI applications.

Question 2

Question Type: MultipleChoice

[InfiniBand Optimization]

Which of the following routing protocols is not capable of avoiding credit loops?

Options:

A- UPDOWN
B- All routing protocols are capable of avoiding credit loops
C- MINHOP
D- FAT TREE

Answer:

C

Explanation:

The MINHOP routing protocol, while efficient in finding minimal paths, does not inherently prevent credit loops. This can lead to deadlocks in the network. In contrast, routing protocols like UPDOWN and FAT TREE are designed to avoid such loops, ensuring more reliable network operation.

Question 3

Question Type: MultipleChoice

[Spectrum-X Configuration]

You are planning to deploy a large-scale Spectrum-X network for AI workloads. Before physical implementation, you want to validate the network design and configuration using a digital twin approach.

Which NVIDIA tool would be most appropriate for creating and simulating a digital twin of your Spectrum-X network?

Options:

A- NVIDIA Base Command Manager
B- NVIDIA Omniverse
C- NVIDIA NetQ
D- NVIDIA Air

Answer:

D

Explanation:

NVIDIA Air is a cloud-based network simulation tool designed to create digital twins of data center infrastructure, including Spectrum-X networks. It allows users to model switches, SuperNICs, and storage components, enabling the simulation, validation, and automation of network configurations before physical deployment. This facilitates Day 0, 1, and 2 operations, ensuring that network designs are tested and optimized for AI workloads.

Reference Extracts from NVIDIA Documentation:

'NVIDIA Air enables cloud-scale efficiency by creating identical replicas of real-world data center infrastructure deployments.'

'NVIDIA Air allows users to model data center deployments with full software functionality, creating a digital twin. Transform and accelerate time to AI by simulating, validating, and automating changes and updates.'

'NVIDIA Air supports simulation of NVIDIA Spectrum Ethernet (Cumulus Linux and SONiC) switches and NVIDIA BlueField DPUs and SuperNICs as well as the NetQ network operations toolset.'

Question 4

Question Type: MultipleChoice

[Spectrum-X Configuration]

You are troubleshooting a Spectrum-X network and need to validate the fabric configuration. Which feature of Spectrum-X allows for automated fabric validation?

Options:

A- NVIDIA NetQ
B- RoCE Adaptive Routing
C- NVIDIA DOCA
D- RoCE Performance Isolation

Answer:

A

Explanation:

NVIDIA NetQ is a network operations tool that provides real-time visibility and automated validation of the network fabric. It helps in identifying misconfigurations, monitoring network health, and ensuring that the fabric meets the required specifications for AI workloads.

Question 5

Question Type: MultipleChoice

[InfiniBand Configuration]

When designing a multi-tenancy East/West (E/W) fabric using Unified Fabric Manager (UFM), which method should be used?

Options:

A- Partition / PKey
B- VLAN
C- ROMA
D- VXLAN

Answer:

A

Explanation:

In InfiniBand networks, Partitioning using Partition Keys (PKeys) is the standard method for implementing multi-tenancy and traffic isolation. PKeys allow administrators to define logical partitions within the fabric, ensuring that traffic is confined to designated groups of nodes. This mechanism is essential for creating secure and isolated environments in multi-tenant architectures.

The Unified Fabric Manager (UFM) leverages PKeys to manage these partitions effectively, enabling administrators to assign and control access rights across different tenants. This approach ensures that each tenant's traffic remains isolated, maintaining both security and performance integrity within the shared fabric.

Question 6

Question Type: MultipleChoice

[InfiniBand Configuration -- Multi-Tenancy with PKey]

You are tasked with configuring multi-tenancy using partition key (PKey) for a high-performance storage fabric running on InfiniBand. Each tenant's GPU server is allowed to access the shared storage system but cannot communicate with another tenant's GPU server.

Which of the following partition key membership configurations would you implement to set up multi-tenancy in this environment?

Options:

A- Assign full membership to both GPU servers and storage system.
B- Assign limited membership to both GPU servers and storage system.
C- Assign limited membership PKey to the shared storage system and full membership PKey to each tenant's GPU servers.
D- Assign full membership PKey to the shared storage system and limited membership PKey to each tenant's GPU servers.

Answer:

D

Explanation:

To enforce strict multi-tenancy, where:

Tenant A's GPU cannot talk to Tenant B's GPU

But both can access shared storage

The correct solution is:

Storage system Full PKey membership

Each tenant's GPU Limited PKey membership

From the NVIDIA InfiniBand P_Key Partitioning Guide:

'A port with limited membership can only communicate with full members of the same PKey. It cannot communicate with other limited members, even within the same partition.'

This isolates tenants from each other, while allowing shared access to storage.

Incorrect Options:

A permits tenant-to-tenant communication.

B isolates everything, including access to storage.

C prevents GPU access to storage.

P2P exams