

PRACTICE QUESTIONS

NVIDIA

NCP-AIN

NVIDIA-Certified Professional: AI Networking

SHARED BY [u/Cozy-Seal-1270](#) on aiexamcerts

EDITION Free Set 1

QUESTIONS 10

SHARED October 18, 2026

EXAM NVIDIA NCP-AIN, NVIDIA-Certified Professional: AI Networking

Shared by u/Cozy-Seal-1270 on aiexamcerts

10 practice questions for the NVIDIA NCP-AIN exam, NVIDIA-Certified Professional: AI Networking, each followed by its answer and a short explanation. More free sets and the full practice exam:

aiexamcerts.com/nvidia/ncp-ain. Spotted a mistake? Email support@aiexamcerts.com with the exam code NCP-AIN and the question number.

1. Which command family in NVUE shows the current state of switch interfaces on a Cumulus Linux switch?

- A vtysh show running-config
- B ethtool -S eth0
- C nv show interface
- D ipmitool sensor list

Answer: C

NVUE uses the nv command, and nv show interface reports the state of switch ports. vtysh belongs to the FRR routing suite and is a tempting distractor for anyone used to older Cumulus or Linux routing CLIs.

2. NVIDIA's term for an advanced network adapter with programmable packet processing and congestion control for AI fabrics is the _____.

Answer: SuperNIC (also accepted: Super NIC)

SuperNIC is the NVIDIA term for ConnectX-class adapters built for AI networks, with advanced packet processing and congestion control. The product term is the one the exam uses, so a generic NIC label would not match.

3. In a BGP-EVPN fabric carrying VXLAN, which VNI type typically carries routed traffic between subnets for a tenant VRF under symmetric IRB?

- A L2VNI
- B VLAN ID 1
- C Management VRF
- D L3VNI

Answer: D

Symmetric IRB routes between tenant subnets using an L3VNI per VRF, while L2VNIs only extend a layer 2 segment. An L2VNI is the tempting distractor because it is the VNI type most people configure first.

4. What does a VRF provide in a multi-tenant Spectrum-X design?
- A A separate routing table per tenant, so overlapping IP ranges stay isolated
 - B Hardware encryption of every packet between leaf and spine switch ports for all tenants
 - C Aggregation of uplinks into a single LAG
 - D Priority queuing for RoCE traffic

Answer: A

A VRF gives each tenant its own routing table, which keeps overlapping address space and routes separated on shared hardware. Priority queuing is a QoS function, which the tempting distractor mixes up with tenant isolation.

5. NVIDIA DSX Air can simulate a network topology virtually so that configuration changes are tested before they reach production switches.
- A True
 - B False

Answer: A (True)

DSX Air provides a simulated network environment for building and testing topologies and configurations. Using it as a pre-production check is the purpose the source describes.

6. A tenant reports intermittent loss on a RoCE training job. Which first step best shows whether the switch is dropping the tenant's traffic?
- A Reboot the leaf switch during the job
 - B Check WJH drop counters for the affected ports
 - C Disable PFC on every port of the fabric
 - D Raise the MTU on the out-of-band management interface

Answer: B

WJH drop reasons on the affected ports show whether the switch discarded the traffic and why, before anything is changed. Disabling PFC would create more loss on RoCE traffic, which makes it a tempting but harmful choice.

7. Which two features help a Spectrum-X fabric respond to uneven load or congestion? (Choose two.)
- A Adaptive routing across equal-cost paths
 - B Spanning Tree blocking of redundant links
 - C Static single-path routing for every flow
 - D Telemetry that shows per-port queue and congestion state
 - E Manual ARP pinning on each server

Answer: A, D

Adaptive routing lets the fabric choose among equal-cost paths based on load, and telemetry shows which ports are congested so operators can act. Spanning Tree blocking removes redundant paths, which works against load spreading in a leaf-spine fabric.

8. Adaptive routing chooses among equal-cost paths using current link load or _____ rather than a fixed hash of the flow.

Answer: congestion (also accepted: queue occupancy, congestion state, queue depth)

Adaptive routing reacts to live congestion signals, such as queue occupancy, so that flows move away from busy links. A static hash keeps a flow on one path even when that path is saturated.

9. Which NVIDIA tool provides real-time monitoring of a Cumulus-based fabric, including congestion detection and latency measurements?
- A NVIDIA Magnum IO
 - B NVIDIA Run:ai
 - C NVIDIA NetQ
 - D NVIDIA Fleet Command

Answer: C

NetQ collects fabric telemetry and reports congestion and latency in real time for Cumulus-based networks. Magnum IO is a communication library for applications, which makes it a tempting but wrong answer for a fabric monitoring question.

10. A leaf is carrying RoCE traffic, but PFC is not enabled on the server-facing ports. Which symptom is most likely during traffic bursts?
- A Routing table overflow on the spine
 - B Fabric manager service stopping
 - C DNS resolution failures on the tenant hosts that need reverse lookups
 - D RoCE flows see packet drops and retransmissions under bursts

Answer: D

RoCE expects a lossless transport, so without PFC bursts can overflow buffers and cause drops that trigger retransmissions and throughput collapse. Routing table overflow is a control-plane symptom that does not follow from a missing lossless class.