NVIDIA.NCP-AIN.v2026-07-10.q30

Exam Code:	NCP-AIN
Exam Name:	NVIDIA-Certified Professional AI Networking
Certification Provider:	NVIDIA
Free Question Number:	30
Version:	v2026-07-10
# of views:	130
# of Questions views:	300
https://www.freepdfdumps.com/NVIDIA.NCP-AIN.v2026-07-10.q30.html	

NEW QUESTION: 1
You are concerned about potential security threats and unexpected downtime in your InfiniBand data center.
Which UFM platform uses analytics to detect security threats, operational issues, and predict network failures in InfiniBand data centers?
A. Host Agent
B. Enterprise Platform
C. Cyber-AI Platform
D. Telemetry Platform
Answer: C (LEAVE A REPLY)
TheNVIDIA UFM Cyber-AI Platformis specifically designed to enhance security and operational efficiency in InfiniBand data centers. It leverages AI-powered analytics to detect security threats, operational anomalies, and predict potential network failures. By analyzing real-time telemetry data, it identifies abnormal behaviors and performance degradation, enabling proactive maintenance and threat mitigation.
This platform integrates with existing UFM Enterprise and Telemetry services to provide a comprehensive view of the network's health and security posture. It utilizes machine learning algorithms to establish baselines for normal operations and detect deviations that may indicate security breaches or hardware issues.
Reference:NVIDIA UFM Cyber-AI Documentation v2.9.1

NEW QUESTION: 2
You have recently implemented NVIDIA Spectrum-X in your data center to optimize AI workloads. You need to verify the performance improvements and create a baseline for future comparisons.
Which tool would be most appropriate for creating performance baseline results in this Spectrum-X environment?

A. NetQ
B. CloudAI Benchmark
C. MLNX-OS
D. Ansible
Answer: B (LEAVE A REPLY)
TheCloudAI Benchmarkis designed to evaluate and establish performance baselines in AI-optimized networking environments like NVIDIA Spectrum-X. It assesses various performance metrics, including throughput and latency, ensuring that the network meets the demands of AI workloads. This benchmarking is essential for validating the benefits of Spectrum-X and for ongoing performance monitoring.
Reference:NVIDIA Spectrum-X Validated Solution Stack

NEW QUESTION: 3
As the network administrator for a large-scale AI research cluster, you are responsible for ensuring seamless data flow across an InfiniBand east-west fabric that interconnects hundreds of compute nodes.
Which tool would you use to trace and discover the network paths between nodes on this InfiniBand east-west fabric?
A. NetQ
B. ibpathverify
C. ibnetdiscover
D. tracert
Answer: C (LEAVE A REPLY)
The ibnetdiscover utility is used to perform InfiniBand subnet discovery and outputs a human-readable topology file. GUIDs, node types, and port numbers are displayed, as well as port LIDs and node descriptions.
All nodes and links are displayed, providing a full topology. This utility can also be used to list the current connected nodes. The output is printed to the standard output unless a topology file is specified.
InfiniBand is a high-performance, low-latency interconnect technology used in AI and HPC data centers, particularly for east-west traffic between compute nodes in large-scale fabrics. Ensuring seamless data flow requires tools to troubleshoot and monitor the network, including the ability to trace and discover network paths between nodes. The question asks for the specific tool used to trace and discover paths in an InfiniBand fabric, which is a key task in InfiniBand troubleshooting.
According to NVIDIA's official InfiniBand documentation, the ibnetdiscover tool is designed to discover and map the topology of an InfiniBand fabric, including the paths between nodes. It scans the fabric, queries the subnet manager, and generates a topology map that details the connections between switches, Host Channel Adapters (HCAs), and other devices. This tool is essential for verifying connectivity, identifying routing paths, and

troubleshooting issues like misconfigured routes or link failures in large-scale InfiniBand fabrics.
Exact Extract from NVIDIA Documentation:
"The ibnetdiscover tool is used to discover the InfiniBand fabric topology and generate a map of the network.
It queries the subnet manager to retrieve information about all nodes, switches, and links in the fabric, providing a detailed view of the paths between nodes. This tool is critical for troubleshooting connectivity issues and ensuring proper routing in InfiniBand networks."
- NVIDIA InfiniBand Networking Guide
This extract confirms that ibnetdiscover is the correct tool for discovering network paths in an InfiniBand east- west fabric. It provides a comprehensive view of the fabric's topology, enabling administrators to trace paths between compute nodes and ensure seamless data flow.
Reference: InfiniBand Fabric Utilities - NVIDIA Docs

NEW QUESTION: 4
What are the necessary steps to upgrade the MLNX-OS on InfiniBand Switches?
A. Connect to the switches using SSH, fetch the MLNX-OS software image, and use the 'install' command to perform the upgrade.
B. Power off the switches, insert the installation media, and power on the switches to start the upgrade process.
C. Restart the switches, connect to the switches using Telnet, and use the 'update' command to perform the upgrade.
D. Remove the switches from the switch fabric, fetch the MLNX-OS software image, and use the 'upgrade' command to perform the upgrade.
Answer: (SHOW ANSWER)
To upgrade the MLNX-OS on InfiniBand switches, the recommended procedure is as follows:
* Connect to the switch via SSH: Establish a secure shell connection to the switch using its management IP address.
* Fetch the MLNX-OS software image: Obtain the appropriate MLNX-OS software image from the official source or repository.
* Use the 'install' command to perform the upgrade: Execute the 'install' command on the switch to initiate the upgrade process with the fetched software image.
This method ensures a smooth and efficient upgrade without the need for physical intervention or service disruption.
Reference Extracts from NVIDIA Documentation:
* "Click on Systems # MLNX-OS Upgrade. Select the desired upgrade method (e.g. 'Install from local file'). Select your image and click 'Install Image'."

NEW QUESTION: 5

In Cumulus Linux, which technology enables the ability to provide active-active redundancy to servers, without the need for direct inter-switch links?

A. MLAG

B. VSS

C. EVPN Multi-homing

Answer: (SHOW ANSWER)

EVPN Multi-homingenablesactive-active redundancy without inter-switch linksby usingoverlay routing over VXLAN and distributed control plane using BGP EVPN.

From the officialNVIDIA Cumulus Linux EVPN Multihoming Documentation:

"EVPN multihoming allows multiple Top-of-Rack (ToR) switches to connect to the same server while maintaining full layer-2 redundancy without the need for inter-switch links or traditional MLAG configuration." Key benefits:

* Simplified topology (no ISL/peer-link needed)
* BGP-based control plane
* Fast convergence
* Active-active links per host NIC

Incorrect Options:

* MLAGrequires ISL between switches and peer-link configuration.
* VSS(Virtual Switching System) is a Cisco term, not supported in NVIDIA networking.

Reference: Cumulus Linux Docs - EVPN Multihoming

NEW QUESTION: 6

You are deploying a Kubernetes cluster for AI workloads using NVIDIA Spectrum-X switches. You need to automate the deployment and management of networking components in this environment.

Which NVIDIA tool is specifically designed to automate the deployment and management of networking components in a Kubernetes cluster with Spectrum-X switches?

A. GPU Operator

B. Network Operator

C. Container Runtime

D. Mellanox OFED

Answer: B (LEAVE A REPLY)

NEW QUESTION: 7

What are two methods for accessing the operating system on a BlueField DPU?

Pick the 2 correct responses below

A. Via the networking interfaces (data ports) in NIC mode

B. Via the rshim interface over the PCIe bus

C. Via the Redfish API

D. Via rshim over a USB connection on the host

Answer: B,D (LEAVE A REPLY)

Accessing the BlueField DPU Operating System (OS) is possible through rshim , either over PCIe or USB, and via SSH through the OOB interface when in DPU mode.
From the NVIDIA BlueField Software Documentation :
"You can access the BlueField OS through the rshim interface. The rshim module enables host-to-DPU communication either via PCIe (default) or USB."
* B. rshim over PCIe : Default when BlueField is installed in a host.
* D. rshim over USB : Useful for provisioning or systems without PCIe drivers.
Incorrect Options:
* A (NIC mode) : BlueField acts as a transparent NIC; OS access is not available to the host.
* C (Redfish) : Redfish is for out-of-band management, not direct OS-level access.
Reference: Accessing BlueField OS - rshim via PCIe and USB Methods

NEW QUESTION: 8
Which of the following options correctly describes the difference between UFM Telemetry, UFM Enterprise, and UFM Cyber AI?
A. UFM Telemetry provides real-time monitoring and analysis of network performance, UFM Enterprise focuses on network management and optimization, and UFM Cyber AI detects and mitigates network security threats.
B. UFM Telemetry provides real-time monitoring and analysis of network performance. UFM Enterprise detects and mitigates network security threats, and UFM Cyber AI focuses on network management and optimization.
C. UFM Telemetry detects and mitigates network security threats. UFM Enterprise provides real-time monitoring and analysis of network performance, and UFM Cyber AI focuses on network management and optimization.
D. UFM Telemetry focuses on network management and optimization, UFM Enterprise detects and mitigates network security threats, and UFM Cyber AI provides real-time monitoring and analysis of network performance.
Answer: A (LEAVE A REPLY)
* UFM Telemetry : Provides real-time monitoring and analysis of network performance, collecting data such as port counters and cable information to assess the health and efficiency of the network.
* UFM Enterprise : Focuses on comprehensive network management and optimization, enabling administrators to monitor, operate, and optimize InfiniBand scale-out computing environments effectively.
* UFM Cyber AI : Detects and mitigates network security threats by analyzing telemetry data to identify anomalies and potential security issues within the network infrastructure.
Reference Extracts from NVIDIA Documentation:
* "UFM Telemetry provides real-time monitoring and analysis of network performance."
* "UFM Enterprise is a powerful platform for managing InfiniBand scale-out computing environments."

* "UFM Cyber-AI enhances the benefits of UFM Telemetry and UFM Enterprise services by detecting and mitigating network security threats."

NEW QUESTION: 9
A cloud service provider is deploying the NVIDIA Spectrum-X Ethernet platform in a multi-tenant environment. To ensure the security and isolation of each tenant's AI workload, the provider wants to implement a feature that prevents unauthorized accessto the network. Which of the following features of the Spectrum-X platform should the provider implement?
A. Streaming Telemetry
B. Adaptive Routing
C. Congestion Control
D. Traffic Isolation
Answer: D (LEAVE A REPLY)
In multi-tenant AI cloud environments, ensuring that each tenant's workloads are isolated and secure is paramount. The NVIDIA Spectrum-X platform addresses this need through itsTraffic Isolationcapabilities.
This feature ensures that network resources are partitioned effectively, preventing unauthorized access and interference between tenants. By implementing Traffic Isolation, the provider can maintain strict boundaries between different tenant environments, ensuring both security and performance consistency.
Reference Extracts from NVIDIA Documentation:
* "Spectrum-X enhances multi-tenancy with performance isolation to ensure tenants' AI workloads perform optimally and consistently."
* "Spectrum-X utilizes the programmable congestion control function on the BlueField-3 hardware platform to accurately assess the congestion condition of the traffic path by using in-band telemetry information... to achieve the goal of performance isolation to ensure that each tenant gets the best expected performance in the cloud and is not negatively affected by congestion of other tenants."

NEW QUESTION: 10
You are troubleshooting a Spectrum-X network and need to ensure that the network remains operational in case of a link failure. Which feature of Spectrum-X ensures that the fabric continues to deliver high performance even if there is a link failure?
A. RoCE Congestion Control
B. RoCE Adaptive Routing
C. NVIDIA NetQ
D. RoCE Performance Isolation
Answer: B (LEAVE A REPLY)
RoCE Adaptive Routing is a key feature of NVIDIA Spectrum-X that ensures high performance and resiliency in the network, even in the event of a link failure. This technology dynamically reroutes traffic to the least congested and operational paths,

effectively mitigating the impact of link failures. By continuously evaluating the network's egress queue loads and receiving status notifications from neighboring switches, Spectrum-X can adaptively select optimal paths for data transmission. This ensures that the network maintains high throughput and low latency, crucial for AI workloads, even when certain links are down.
Reference Extracts from NVIDIA Documentation:
* "Spectrum-X employs global adaptive routing to quickly reroute traffic during link failures, minimizing disruptions and preserving optimal storage fabric utilization."
* "RoCE Adaptive Routing avoids congestion by dynamically routing large AI flows away from congestion points. This approach improves network resource utilization, leaf/spine efficiency, and performance."

NEW QUESTION: 11
You are deploying a Kubernetes cluster for AI workloads using NVIDIA Spectrum-X switches. You need to automate the deployment and management of networking components in this environment.
Which NVIDIA tool is specifically designed to automate the deployment and management of networking components in a Kubernetes cluster with Spectrum-X switches?
A. Mellanox OFED
B. Container Runtime
C. Network Operator
D. GPU Operator
Answer: C (LEAVE A REPLY)
TheNVIDIA Network Operatoris designed to simplify and automate the deployment and management of networking components in Kubernetes clusters, particularly those utilizing NVIDIA Spectrum-X switches. It manages the installation and configuration of necessary drivers, plugins, and other networking resources to enable features like RDMA and GPUDirect RDMA, which are essential for high-performance AI workloads.
By leveraging Kubernetes Custom Resource Definitions (CRDs) and the Operator Framework, the Network Operator ensures that networking components are consistently and correctly configured across the cluster, reducing manual intervention and potential configuration errors.
Reference:NVIDIA Network Operator Documentation

NEW QUESTION: 12
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?

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: (SHOW ANSWER)
To enforce strictmulti-tenancy, where:
* Tenant A's GPUcannot talk toTenant B's GPU
* But both can accessshared storage
The correct solution is:
* Storage system # Full PKey membership
* Each tenant's GPU # Limited PKey membership
From theNVIDIA 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 tenantsfrom each other, while allowingshared access to storage.
Incorrect Options:
* Apermits tenant-to-tenant communication.
* Bisolates everything, including access to storage.
* Cprevents GPU access to storage.
Reference: NVIDIA InfiniBand - Multi-Tenant PKey Partitioning Design

NEW QUESTION: 13
Which of the following tools in Cumulus Linux is specifically useful for detecting and differentiating microbursts from regular network congestion?
Pick the 2 correct responses below
A. Monthly network utilization reports
B. ASIC monitoring with millisecond-level granularity
C. SNMP polling at 5-minute intervals
D. What Just Happened (WJH) feature for packet drop analysis
Answer: B,D (LEAVE A REPLY)
In Cumulus Linux, microbursts are short-lived, high-volume traffic bursts that often go undetected by coarse- grained monitoring like SNMP.
The two tools specifically used for this purpose are:
* What Just Happened (WJH)
"WJH provides real-time packet drop visibility and classifies drops by reason (e.g., congestion, ACLs, etc.), enabling microburst detection."
* ASIC monitoring at millisecond granularity
"Deep telemetry is enabled via the switch ASIC, which provides sub-second counters that capture microburst patterns otherwise missed by SNMP." Incorrect Options:

* AandCprovide low-frequency sampling, insufficient for microbursts which last milliseconds.
Reference: NVIDIA NetQ & Cumulus Linux Documentation - What Just Happened (WJH)

NEW QUESTION: 14
What is the purpose of configuring NVUE to ignore Linux files?
A. Enable pushing of configuration through Ansible template files.
B. Enable the persistent manipulation of specific settings using both NVUE and flat-file approaches.
C. Reduce NVUE memory utilization to optimize performance.
D. Improve Cumulus security by reducing the attack surface.
Answer: B (LEAVE A REPLY)
Configuring NVUE to ignore certain underlying Linux files allows administrators to manage specific settings manually or through automation tools like Ansible withoutNVUE overwriting these configurations. This approach enables the persistent manipulation of settings using both NVUE and flat-file methods, providing flexibility in network management.

NEW QUESTION: 15
What is the basic functionality of an IB Router?
A. Connecting Ethernet switches to an InfiniBand fabric
B. Connecting SAN to an InfiniBand fabric
C. Connecting NVLink domains to an InfiniBand fabric
D. Connecting two (or more) InfiniBand fabrics
Answer: (SHOW ANSWER)
AnInfiniBand (IB) Routerconnects two or moreInfiniBand subnets, making it possible for nodes in different subnets to communicate throughroute-managed communication.
From the officialNVIDIA InfiniBand Routers Documentation:
"An InfiniBand router provides connectivity between two or more InfiniBand subnets, enabling communication between hosts that are not on the same subnet while preserving isolation and scalability."
* Ensures fabric scalability by allowing subnet segmentation.
* Uses LID routing across subnet managers (SMs).
* Essential in large clusters with thousands of nodes.
Incorrect Options:
* AandBare incorrect: InfiniBand does not connect directly to Ethernet or SANs without a gateway.
* Cis unrelated: NVLink is a GPU interconnect, not tied to InfiniBand routers.
Reference: NVIDIA InfiniBand Routers Guide

NEW QUESTION: 16

You are troubleshooting connectivity issues in your InfiniBand network and need to test basic connectivity between nodes. Which command should you use to test basic connectivity between InfiniBand nodes?

A. ibping

B. traceroute

C. ibnetdiscover

D. ping

Answer: A (LEAVE A REPLY)

The tool specifically designed for testingInfiniBand connectivityis **ibping**. It functions similarly to the traditional ping utility but is optimized forInfiniBand fabrics.

From theNVIDIA InfiniBand Diagnostic Utilities Documentation:

"ibping tests the connectivity of InfiniBand nodes by sending management datagrams (MADs) and verifying the response from the destination LID or GUID."

* Tests basicnode-to-nodereachability
* Supports testing viaLID, GUID, orport number
* Helps verify subnet manager routing and fabric health

Incorrect Options:

* pingandtracerouteare IP-based, not fabric-aware.
* ibnetdiscovermaps topology but doesn't test live connectivity.

Reference: InfiniBand Diagnostic Tools - ibping

NEW QUESTION: 17

In order to configure RoCE on a Cumulus switch, which command should be used?

A. nv set qos roce enable on

B. nv set roce qos enable on

C. nv roce qos enable on

D. nv qos roce enable on

Answer: A (LEAVE A REPLY)

To enable RDMA over Converged Ethernet (RoCE) on a Cumulus switch, the correct command is:

nv set qos roce enable on

This command configures the Quality of Service (QoS) settings to support RoCE, ensuring that the necessary parameters for lossless Ethernet are applied.
Reference: NVIDIA Cumulus Linux Documentation - RDMA over Converged Ethernet (RoCE)

NEW QUESTION: 18
Which component of the Spectrum-X platform is responsible for reordering out-of-order packets?
A. SuperNIC
B. Spectrum-4 switch
C. DOCA software
D. NetQ
Answer: (SHOW ANSWER)
Within the Spectrum-X platform, the NVIDIA BlueField-3 SuperNIC is responsible for reordering out-of- order packets. When RoCE adaptive routing is employed, packets may arrive at their destination out of order due to dynamic path selection. The BlueField-3 SuperNIC handles this by reassembling the packets in the correct order at the transport layer, ensuring that the application receives data seamlessly.
Reference Extracts from NVIDIA Documentation:
* "As different packets of the same flow travel through different paths of the network, they may arrive out of order to their destination. At the RoCE transport layer, the BlueField-3 DPU takes care of the out- of-order packets and forwards the data to the application in order."
* "The BlueField-3 SuperNIC offers adaptive routing, out-of-order packet handling and optimized congestion control." The NVIDIA Spectrum-X networking platform is an Ethernet-based solution optimized for AI workloads, combining Spectrum-4 switches, BlueField-3 SuperNICs, and software like DOCA and NetQ to deliver high performance, low latency, and efficient data transfer. A key feature of Spectrum-X is its adaptive routing, which dynamically selects the least-congested paths for packet transmission to maximize bandwidth and minimize latency. However, this per-packet load balancing can result in packets arriving out of order at the destination, necessitating a mechanism to reorder them for seamless application performance. The question asks which Spectrum-X component is responsible for reordering these out-of-order packets.
According to NVIDIA's official documentation, the BlueField-3 SuperNIC is the component responsible for reordering out-of-order packets in the Spectrum-X platform. The SuperNIC, a network accelerator designed for hyperscale AI workloads, handles packet reordering at the RDMA over Converged Ethernet (RoCE) transport layer. It uses its processing capabilities to transparently reorder packets and place them in the correct sequence in the host memory, ensuring that adaptive routing's out-of-order delivery is invisible to the application. This is critical for maintaining predictable performance in AI workloads, particularly for GPU-to- GPU communication in Spectrum-X networks.

Exact Extract from NVIDIA Documentation:
"The Spectrum-4 switches are responsible for selecting the least-congested port for data transmission on a per- packet basis. As different packets of the same flow travel through different paths of the network, they may arrive out of order to their destination. The BlueField-3 SuperNIC transforms any out-of-order data at the RoCE transport layer, transparently delivering in-order data to the application."
- NVIDIA Technical Blog: Turbocharging Generative AI Workloads with NVIDIA Spectrum-X Networking Platform This extract confirms that option A, the SuperNIC (specifically the BlueField-3 SuperNIC), is the correct answer. The SuperNIC's role in reordering packets ensures that the adaptive routing implemented by Spectrum-4 switches does not compromise application performance, maintaining high effective bandwidth and low tail latency for AI workloads.

NEW QUESTION: 19
You are configuring the Unified Fabric Manager (UFM) for an InfiniBand fabric in a multi-tenant environment. You need to implement a solution that can detect potential security threats.
Which UFM feature uses analytics to detect security threats and predict network failures in InfiniBand data centers?
A. Host Agent
B. Telemetry platform
C. Cyber-AI platform
D. Enterprise platform
Answer: C (LEAVE A REPLY)
The UFM Cyber-AI platform is an advanced feature of NVIDIA's Unified Fabric Manager designed to enhance security and reliability in InfiniBand data centers. It leverages AI-powered analytics and machine learning techniques to detect security threats, operational anomalies, and predict potential network failures.
By analyzing real-time and historical telemetry data, UFM Cyber-AI can identify abnormal system behaviors, performance degradations, and usage profile changes. This proactive approach enables administrators to address issues before they escalate, ensuring the integrity and uptime of the data center.
Reference Extracts from NVIDIA Documentation:
* "The NVIDIA Unified Fabric Manager (UFM) Cyber-AI platform offers enhanced and real-time network telemetry, combined with AI-powered intelligence and advanced analytics. It enables IT managers to discover operational anomalies and even predict network failures."
* "UFM Cyber-AI uses machine learning (ML) techniques and AI models for anomaly detection and prediction to learn the lifecycle patterns of data center network components."
* "The NVIDIA UFM platforms revolutionize data center networking management by combining enhanced, real-time network telemetry with AI-powered cyber intelligence and analytics to support scale-out InfiniBand data centers. ... The UFM Cyber-AI platform takes

fabric management to the next level by adding an analytics layer powered by artificial intelligence. It enables data center operators to proactively monitor and manage the InfiniBand fabric, predicting and preventing potential failures, optimizing performance, and enhancing security. By analyzing telemetry data and historical patterns, UFM Cyber-AI can detect anomalies that may indicate security threats or operational issues, providing actionable insights to prevent downtime."

NEW QUESTION: 20
A cloud service provider is deploying the NVIDIA Spectrum-X Ethernet platform in a multi-tenant environment. To ensure the security and isolation of each tenant's AI workload, the provider wants to implement a feature that prevents unauthorized access to the network.
Which of the following features of the Spectrum-X platform should the provider implement?
A. Adaptive Routing
B. Traffic Isolation
C. Streaming Telemetry
D. Congestion Control
Answer: B (LEAVE A REPLY)

NEW QUESTION: 21
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?
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 (LEAVE A REPLY)
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.

NEW QUESTION: 22
What does NetQ leverage (in addition to NVIDIA "What Just Happened" switch telemetry data and NVIDIA DOCA telemetry) to help network operators proactively identify server and application root cause issues?
A. Flow telemetry
B. Behavioral telemetry
C. Application telemetry

D. Packet capture telemetry

Answer: B (LEAVE A REPLY)

NetQintegrates multiple telemetry sources, includingWJH,DOCA, and notably,Behavioral Telemetry.

From theNetQ Documentation - Behavioral Telemetry Section:

"Behavioral telemetry in NetQ correlates server and application behavior with network events, offering insights into root cause analysis by detecting anomalies in protocol, path, or performance behavior." This helps identify patterns like:

* Misbehaving applications causing retransmits.
* Sudden changes in traffic flows.
* Latency spikes correlated with app-level issues.

It complements device-level telemetry by introducingintent-based anomaly detection, crucial for proactive operations.

Incorrect Options:

* Flow telemetryandpacket captureoffer raw data but not behavioral insights.
* Application telemetryis too vague and is not the term NetQ uses for this feature.

Reference: NetQ 3.2 Documentation - Behavioral Telemetry

NEW QUESTION: 23

Which command should be used to configure RoCE on a Cumulus switch?

A. nv set roce qos enable on

B. nv set qos roce enable on

C. nv qos roce enable on

D. nv roce qos enable on

Answer: B (LEAVE A REPLY)

Comprehensive and Detailed 150 to 250 words of Explanation From NVIDIA AI Networking Topics:

The correct answer is B. nv set qos roce enable on . In NVIDIA Cumulus Linux, RoCE configuration is managed under the QoS hierarchy in NVUE because RoCE depends on lossless Ethernet behavior, priority flow control, congestion management, and related quality-of-service mechanisms. NVIDIA's official Cumulus Linux documentation shows that the RoCE feature is configured through nv set qos roce , and in newer command references this appears explicitly as nv set qos roce enable on|off . This confirms both the command structure and the ordering of the keywords: nv set # qos # roce # enable on .

The other options are incorrect because they do not follow valid NVUE syntax. NVIDIA documents NVUE commands in a strict object-based order, where the top-level feature group comes immediately after nv set .

Since RoCE is implemented as part of the QoS framework, qos must appear before roce . Commands such as nv set roce qos enable on , nv qos roce enable on , and nv roce qos enable on do not match the official NVIDIA command model. Therefore, for enabling RoCE

on a Cumulus switch in an NVIDIA AI networking environment, the verified command is nv set qos roce enable on .

NEW QUESTION: 24
You are optimizing an AI workload that involves multiple GPUs across different nodes in a data center. The application requires both high-bandwidth GPU-to-GPU communication within nodes and efficient communication between nodes.
Which combination of NVIDIA technologies would best support this multi-node, multi-GPU AI workload?
A. NVLink for both intra-node and inter-node GPU communication.
B. InfiniBand for both intra-node and inter-node GPU communication.
C. NVLink for intra-node GPU communication and InfiniBand for inter-node communication.
D. PCIe for intra-node GPU communication and RoCE for inter-node communication.
Answer: C (LEAVE A REPLY)
For optimal performance in multi-node, multi-GPU AI workloads:
* NVLinkprovides high-speed, low-latency communication between GPUs within the same node.
* InfiniBandoffers efficient, scalable communication between nodes in a data center.Combining these technologies ensures both intra-node and inter-node communication needs are effectively met.
Reference:NVIDIA NVLink & NVSwitch: Fastest HPC Data Center Platform

NEW QUESTION: 25
What are the prerequisites for performing Flow Analysis with NetQ?
A. Cumulus 4.x and later / Spectrum-2 and later / LCM enabled
B. Cumulus 5.x and later / Spectrum-3 and later / On-premises deployment
C. Cumulus 5.x and later / Spectrum-2 and later / On-premises deployment
D. Cumulus 5.x and later / Spectrum-2 and later / LCM enabled
Answer: (SHOW ANSWER)
To perform Flow Analysis with NetQ, the following prerequisites must be met:
* Cumulus Linux Version: NetQ Flow Analysis requires Cumulus Linux 5.x or later.
* Switch Hardware: The feature is supported on Spectrum-2 and later switch models.
* Lifecycle Management (LCM): LCM must be enabled to utilize Flow Analysis capabilities.
These requirements ensure compatibility and proper functioning of the Flow Analysis feature within NetQ.
Reference: NVIDIA NetQ Documentation - Flow Analysis Prerequisites

NEW QUESTION: 26
A leading AI research center is upgrading its infrastructure to support large language model projects.

The team is debating whether to implement a dedicated storage fabric for their AI workloads.
Which of the following best explains why a dedicated storage fabric is crucial for this AI network architecture?
Pick the 2 correct responses below
A. To enable parallel data access and improve storage performance for distributed AI workloads.
B. To ensure data security and isolation from other network traffic.
C. To provide high-bandwidth, low-latency data access that prevents I/O bottlenecks during AI model training.
D. To reduce the overall cost of the storage infrastructure.
Answer: A,C (LEAVE A REPLY)
Modern AI training (especially with LLMs) requires extremely high-speed, parallel access to large datasets. A dedicated storage fabric separates data I/O traffic from the training compute path and avoids contention.
From NVIDIA DGX Infrastructure Reference Architectures :
"Dedicated storage networks eliminate I/O bottlenecks by providing low-latency, high-bandwidth access to distributed storage for large-scale training jobs."
"Parallel access to datasets is key for performance, especially in multi-node, multi-GPU AI clusters." Security (B) is important, but not the core reason for a storage fabric.
Cost (D) is typically increased , not reduced, with dedicated fabrics.
Reference: NVIDIA BasePOD/AI Infrastructure Deployment Guidelines - Storage Section

NEW QUESTION: 27
What are the two general user account types in MLNX-OS?
Pick the 2 correct responses below:
A. viewer
B. monitor
C. admin
D. enable
Answer: B,C (LEAVE A REPLY)
MLNX-OS, the operating system for NVIDIA's networking devices, defines two primary user account types:
adminandmonitor. Theadminaccount has full administrative privileges, allowing for complete configuration and management of the system. Themonitoraccount, on the other hand, is designed for users who need to view system configurations and statuses without making any changes. This separation ensures a clear distinction between users who manage the system and those who monitor its operations.
Reference Extracts from NVIDIA Documentation:
* "There are two user roles or account types: admin and monitor. As 'admin', the user is privileged to run all the available commands. As 'monitor', the user can run commands that

show system configuration and status, or set terminal settings." MLNX-OS is the network operating system used on NVIDIA's Mellanox Ethernet switches, including the Spectrum family (e.g., Spectrum-4 switches in the Spectrum-X platform), designed for high-performance Ethernet networking in AI and HPC data centers. MLNX-OS provides a command-line interface (CLI) for configuring and managing switch operations, with user accounts controlling access to various commands and functions. The question asks for the two general user account types in MLNX-OS, which define the primary privilege levels for user access.
According to NVIDIA's official MLNX-OS documentation, the two general user account types in MLNX-OS are:
* monitor: This account type has read-only access, allowing users to view configurations, status, and logs but not modify settings. It is used for monitoring and troubleshooting without risking unintended changes.
* admin: This account type has full read-write access, enabling users to view and modify all configurations, execute commands, and manage the switch's operations. It is intended for administrators with complete control over the system.
These two account types represent the primary privilege levels in MLNX-OS, providing a clear distinction between read-only monitoring and full administrative access.
Exact Extract from NVIDIA Documentation:
"MLNX-OS supports two primary user account types for managing switch operations:
* monitor: Users with monitor privileges have read-only access to the system. They can view configuration details, system status, and logs but cannot make changes to the configuration.
* admin: Users with admin privileges have full read-write access, allowing them to configure, manage, and troubleshoot all aspects of the switch, including executing privileged commands.These account types ensure secure and controlled access to the switch's management functions."-NVIDIA MLNX-OS User Manual This extract confirms that options B (monitor) and C (admin) are the correct answers. These account types are the standard privilege levels in MLNX-OS, used to manage access for monitoring and administrative tasks on Spectrum switches, including those in Spectrum-X deployments.

NEW QUESTION: 28
You're designing a multi-GPU system for AI training using NVIDIA GPUs with NVLink connections.
You need to maximize inter-GPU communication bandwidth. Which feature included in NCCL allows for improved communication between GPUs and NICs?
A. Adaptive Routing
B. PXN
C. Graph Search Optimization
D. SHARP v2
Answer: B (LEAVE A REPLY)

The correct answer isPXN (Peer eXchange Network).
From theNVIDIA NCCL Documentation:
"PXN enables communication between GPUs connected via NVLink and NICs by treating the GPUs as a distributed switch. This architecture improves bandwidth utilization by enabling any GPU to communicate with the NIC via the shortest path available, even if it's not directly connected to the NIC." This enhances GPU-to-NIC and NIC-to-GPU transfers, leveraging the NVLink topology. It significantly boosts performance in multi-GPU setups where not every GPU is directly connected to the NIC.
Other options:
* Adaptive Routingis a fabric-level feature for dynamic path rerouting.
* Graph Search Optimizationis used internally for topology modeling in NCCL.
* SHARP v2is a switch-based collective acceleration method, unrelated to PXN.
Reference: NVIDIA NCCL User Guide - PXN Feature Section

NEW QUESTION: 29
You are setting up PKey memberships for different tenants in an InfiniBand network. You want to ensure that some tenants have limited communication capabilities. Which PKey membership type allows members to communicate with full members but not with other members of the same type?
A. Restricted membership
B. Isolated membership
C. Full membership
D. Limited/partial membership
Answer: D (LEAVE A REPLY)
In InfiniBand networks,P_Keys (Partition Keys)control communication boundaries. Each port can belong to one or more partitions with eitherfullorlimitedmembership.
From NVIDIA InfiniBand Documentation (Partitioning and P_Keys):
"A limited (or partial) membership permits a port to communicate only with other ports in the same partition that have full membership. It cannot communicate with other limited members, even if they are in the same P_Key partition." This makes limited/partial membership ideal formulti-tenant security, where tenant ports can reach infrastructure ports (full members) butnotother tenant ports (limited members).
Incorrect Options:
* A & Bare not valid InfiniBand P_Key types.
* C (Full membership)allows unrestricted communication within the same partition.
Reference: NVIDIA InfiniBand Guide - PKey Partitioning and Membership Types

NEW QUESTION: 30
What is the purpose of WJH (What Just Happened)?
A. Provide contextual information regarding dropped packets in order to aid debugging.
B. Send notifications of failed login attempts to a pre-defined Slack channel.

C. Identify potential cyberattacks or unusual traffic patterns across the cluster.
D. Collate operating system logs and diagnose system crashes.
Answer: A (LEAVE A REPLY)
NVIDIA's What Just Happened (WJH) is a feature that provides real-time visibility into network problems by analyzing all packets passing through the switch and alerting on performance issues caused by packet drops, congestion, high latency, or misconfigurations.
WJH retains the last packets that were dropped from the switch with complete packet headers and the actual drop reason. This enhances the ability to debug network problems, identify affected flows, and decrease time- to-repair.