DUMPSARENA

NVIDIA AI Infrastructure

NVIDIA NCP-AII

Version Demo

Total Demo Questions: 8

Total Premium Questions: 71

Buy Premium PDF

https://dumpsarena.co

sales@dumpsarena.co

sales@dumpsarena.co
dumpsarena.co

QUESTION NO: 1

A 24-hour HPL burn-in fails with "illegal value" errors during the first iteration. Which initial troubleshooting step resolves this without compromising burn-in validity?

A. Switch from FP64 to FP32 precision.

B. Disable GPU affinity.

C. Reduce test duration to 12 hours.

D. Verify the matrix size is divisible by block size.

ANSWER: D

Explanation:

In HPL, an “illegal value” error very early in the run is most commonly caused by an invalid or inconsistent `HPL.dat` configuration rather than a hardware fault. A key first check is that the problem size (`N`) and the algorithmic blocking (`NB`) are chosen sensibly so the factorization and panel operations can be partitioned cleanly across the process grid (`P×Q`) and the underlying BLAS kernels. Ensuring `N` is compatible with `NB` (a common best practice is choosing `N` as a multiple of `NB`, and also consistent with the process grid) prevents parameter/pathological cases that can trigger immediate argument-check failures (“illegal value”) while keeping the test mathematically equivalent and therefore still a valid burn-in.

Switching to FP32 (A) changes the workload and no longer represents the standard HPL FP64 stress/burn-in used to validate numerical stability. Disabling GPU affinity (B) may affect performance and locality but is not the typical root cause of an “illegal value” at iteration 1. Reducing duration (C) doesn’t address the configuration error and also weakens the burn-in objective. For HPL tuning and parameter constraints, see the HPL documentation and NVIDIA’s HPL guidance: https://www.netlib.org/benchmark/hpl/ and https://docs.nvidia.com/nvidia-hpc-benchmarks/index.html.

QUESTION NO: 2

A system administrator noticed a failure on a DGX H100 server. After a reboot, only the BMC is available. What could be the reason for this behavior?

A. The network card has no link / connection.

B. A boot disk has failed.

C. Multiple GPUs have failed.

D. There are more than two failed power supplies.

ANSWER: B

Explanation:

If, after a reboot, the only thing you can reliably reach is the BMC, it usually means the BMC is healthy (it’s an independent management controller with its own firmware and network path) but the host platform is not completing a normal boot. On DGX H100, the OS boot volume is on the internal NVMe boot drives (commonly configured as a mirrored/RAID-1 boot set). If

the boot volume is corrupted or both members of the mirror are unavailable, the system may fail to find a bootable device or hang early in POST/UEFI, while the BMC remains accessible for out-of-band management. That makes a boot disk failure (or boot volume corruption) the most plausible cause among the options.

Option A is unlikely because a host NIC link issue wouldn't prevent the server from booting locally; it would mainly affect in-band network access. Option C is also unlikely because GPU failures typically don't block the CPU/OS from booting; you'd more often see degraded GPU inventory or driver issues once the OS is up. Option D is not the best fit because significant PSU failure generally leads to power instability or inability to power the host at all; if the system can power-cycle and the BMC is up, total PSU loss is less consistent than a storage/boot issue.

References: NVIDIA DGX H100 User Guide, NVIDIA DGX Documentation

QUESTION NO: 3

You are installing the operating system as part of the initial setup for a new NVIDIA Base Command Manager (BCM) cluster. Which two of the following actions are essential for a successful OS installation on the cluster's head node? (Pick the 2 correct responses below)

A. Configure network switches for PXE boot to all compute nodes before installing the OS on the head node.

B. Download the latest BCM ISO and verify its integrity using the provided checksum, then start the installation.

C. Start the head node OS installation process with the system BIOS set to legacy boot mode instead of UEFI.

D. Set the desired time zone and configure NTP synchronization during the OS installation wizard.

ANSWER: B D

Explanation:

For a successful head node OS install in NVIDIA Base Command Manager (BCM), two practical "must-dos" are: (1) ensure you are installing from a valid, uncorrupted installer image, and (2) ensure the head node has correct time configuration. Verifying the downloaded BCM/Bright installer ISO via checksum is essential because a corrupted ISO commonly leads to unpredictable installer failures, package errors, or post-install issues that are hard to troubleshoot. Correct time zone plus NTP/chrony configuration is also essential because cluster services rely on consistent time for authentication (e.g., Kerberos/LDAP integrations), certificate validity, logging, and reliable coordination between management services and nodes.

Configuring network switches for PXE to compute nodes is important for later provisioning, but it is not essential to successfully install the head node OS itself; PXE/DHCP/TFTP services are typically configured from the head node after the base install. Setting BIOS to legacy boot mode is not a general requirement and is often wrong on modern platforms where UEFI is expected/recommended; forcing legacy can prevent booting or complicate installation depending on hardware and OS settings.

References: NVIDIA Base Command Manager Documentation, RHEL Time Synchronization (chrony/NTP).

QUESTION NO: 4

During cluster validation, the Cable Validation Tool (CVT) reports "Underperforming (BER)" for an InfiniBand link. Which BER thresholds indicate a critical signal quality issue requiring cable replacement?

A. Rx power variance > 3dB between lanes

B. Effective BER > 0 during the first 125 minutes of link operation

C. Raw BER > 1e-12 or Effective BER > 1.5E-254 for <6hr measurements

D. Temperature > 85Â°C on transceiver module

ANSWER: C

Explanation:

In NVIDIA/Mellanox InfiniBand environments, CVT/UFM flag a link as “Underperforming (BER)” when the measured bit error behavior indicates the physical channel (cable/optics) is no longer providing adequate signal integrity. The key distinction is between *raw* (pre-FEC) errors and *effective* (post-FEC) errors. A non-zero raw BER can be tolerated because Forward Error Correction is expected to correct occasional symbol/bit errors at very high line rates. However, once raw BER rises beyond the typical acceptable envelope (commonly around 1e-12), the FEC margin is being consumed and the link becomes unstable under load. More importantly, any non-negligible post-FEC (“effective”) BER indicates errors are escaping correction, which translates into packet loss/retries and severe performance impact for collective-heavy AI workloads.

Option C is the only choice that directly states BER-based thresholds (raw BER > 1e-12 and/or effective BER exceeding an extremely low bound over the measurement window) that align with how CVT/UFM interpret critical signal-quality conditions that warrant remediation such as cable replacement. Options A and D are potential *contributors* (power imbalance, overheating) but are not BER thresholds. Option B is incorrect because “effective BER > 0 in the first 125 minutes” is not a standard critical threshold definition and is overly simplistic without the defined CVT/UFM windowing/limits.

References: NVIDIA UFM Documentation, NVIDIA Networking Documentation

QUESTION NO: 5

You are a network administrator responsible for configuring an East-West (E/W) Spectrum-X fabric using SuperNIC. The Bluefield-3 devices in your network should be set to NIC mode with RoCE enabled to optimize data flow between servers. You have access to the Spectrum-X management tools and the necessary documentation. You need to use specific configuration commands to achieve this setup. Which of the following steps and commands are necessary to configure the Bluefield-3 devices in NIC mode for the E/W Spectrum-X fabric using SuperNIC? (Pick the 2 correct responses below)

A. Use the command sudo mlxconfig -d /dev/mst/ set LINK_TYPE_P1=2 to enable Ethernet on the Bluefield-3 devices.

B. Use the command sudo mlxconfig -d /dev/mst/ set DISABLE_SPECTRUM_X=1 to reduce overhead.

C. Use the command sudo mlxconfig -d /dev/mst/ set INTERNAL_CPU_OFFLOAD_ENGINE=1 to configure the SuperNIC to operate in NIC mode.

D. Use the command sudo mlxconfig -d /dev/mst/ set DPU_MODE=1 to set up the Bluefield-3 devices in DPU mode.

ANSWER: A C

Explanation:

For an East-West Spectrum-X deployment using BlueField-3 as a SuperNIC, two key prerequisites are (1) ensuring the ports are operating as Ethernet (not InfiniBand) and (2) enabling the SuperNIC offload path used for Spectrum-X features (RoCE acceleration and congestion-control related offloads). Option A is correct because setting `LINK_TYPE_P1=2` (and similarly `LINK_TYPE_P2=2` if applicable) configures the port link type to Ethernet, which is required for RoCE over Ethernet fabrics. Option C is also correct because enabling the internal CPU/offload engine is part of configuring BlueField-3 to operate as a

SuperNIC/NIC-mode device for Spectrum-X, allowing the NIC to handle specific acceleration/offload functions while the host remains in control.

Option B is incorrect: there is no standard/best-practice Spectrum-X setup step to “disable Spectrum-X” via `DISABLE_SPECTRUM_X`; disabling would be counter to the goal and the knob is not a commonly documented mlxconfig setting for this purpose. Option D is incorrect because setting DPU mode is the opposite of the requirement (NIC mode). DPU mode shifts more control/processing to the BlueField Arm complex, whereas SuperNIC/NIC mode is intended for host-managed networking with NIC offloads.

References: NVIDIA Networking Documentation, mlxconfig (MFT) documentation.

QUESTION NO: 6

A team is installing the NVIDIA Run:ai control plane on a Kubernetes cluster. Which two (2) options are most critical to validate before proceeding? (Pick the 2 correct responses below)

A. Helm is installed on the installer machine.

B. Ensure Kubernetes is running on the cluster.

C. All cluster nodes have NVIDIA GPUs installed.

D. NTP is disabled to simplify time synchronization.

ANSWER: A B

Explanation:

Before installing the NVIDIA Run:ai control plane, the two most critical validations are that you have a functioning Kubernetes cluster to deploy into and that you can deploy the Run:ai components using the supported installation tooling (commonly Helm). Run:ai’s control plane is delivered as Kubernetes resources (Deployments, Services, etc.), so if Kubernetes isn’t up and reachable (API server accessible, nodes Ready), the installation cannot proceed at all. Likewise, Run:ai installations are typically performed via Helm charts; if Helm isn’t available on the machine performing the install (or in the CI/CD environment), you won’t be able to render/apply the charts in the supported way, and upgrades/rollbacks become problematic.

Having NVIDIA GPUs on *all* nodes is not a prerequisite for installing the control plane—many clusters have a mix of GPU and non-GPU nodes, and the control plane can run on CPU-only nodes while GPU worker nodes are added/registered later. Finally, disabling NTP is the opposite of best practice: Kubernetes and TLS-based systems rely on accurate time for certificate validation, auditing, and log correlation; you should ensure time sync is enabled and consistent, not disabled.

References: Helm documentation (installing Helm), Kubernetes setup documentation.

QUESTION NO: 7

During multi-node HPL burn-in, GPUs show uneven utilization. Which configuration ensures balanced workload distribution?

A. Enable HPL_USE_NVSHMEM=1 for shared memory acceleration

B. HPL_RUN_GEMM_TESTS to skip validation

C. Set --gpu-affinity and --cpu-affinity to align GPU and NUMA nodes

D. HPL_OOC_TILE_M to 8192 for larger blocks

ANSWER: C

Explanation:

Balanced GPU utilization in multi-node HPL (especially on systems with multiple CPU sockets and many GPUs per node) depends heavily on correct CPU/GPU/NUMA affinity. If MPI ranks (or CPU threads driving GPU work) are scheduled on cores from a different NUMA domain than the GPU they control, traffic can traverse inter-socket links (e.g., UPI) and non-local PCIe paths. That adds latency and reduces effective bandwidth, so some GPUs receive work/data later than others, showing up as uneven utilization. Setting explicit `--cpu-affinity` and `--gpu-affinity` (or equivalent binding via the launcher) pins each rank to CPU cores local to the GPU/NUMA node it uses, minimizing cross-socket traffic and producing a much flatter, consistently high utilization profile during burn-in.

Option A (NVSHMEM) can help certain communication patterns, but it does not inherently fix rank-to-GPU locality issues. Option B is unrelated to performance balance (it changes validation/testing behavior). Option D changes tiling/blocking and may affect performance, but it won't correct a topology/NUMA mismatch that causes some GPUs to stall.

References: NVIDIA NCCL User Guide (env vars and affinity-related behavior), Open MPI mpirun documentation (binding/mapping concepts).

QUESTION NO: 8

An engineer needs to validate 400G DAC cable signal integrity in a DGX cluster. Which CVT metric best identifies marginal cables needing replacement?

A. Lane power variance < 3dB across all transceivers.

B. Transceiver model matching QSFP-DD specifications.

C. Temperature fluctuations > 5Â°C during validation.

D. Effective BER > 1.5E-254 during a <6-hour monitoring window.

ANSWER: D

Explanation:

For 400G (e.g., NDR/400G-class) links, the most actionable way to spot marginal copper (DAC) is to look at bit errors as seen by the link's error-correction pipeline. CVT/UFM-style health checks focus on BER counters because DACs don't provide optical diagnostics like Rx/Tx power; the cable's "signal integrity" shows up as symbol/bit errors on the lanes. "Effective BER" (post-FEC) is the key metric because it represents errors that remain after Forward Error Correction—i.e., what can actually impact packet integrity and force retries, congestion control reactions, and ultimately NCCL/all-reduce performance instability. A non-zero Effective BER over a monitoring window is a strong indicator the link is operating too close to the margin and the cable should be investigated/replaced.

Option A is not applicable to DAC validation because "lane power variance" is an optical-module concept (DOM) and not a primary DAC integrity indicator. Option B is a procurement/compatibility check, not a signal-integrity metric. Option C (temperature swings) can correlate with issues but is indirect and not the best CVT metric to identify a marginal cable. Therefore, Effective BER thresholding is the best choice.

References: NVIDIA UFM Documentation, NVIDIA Networking – Forward Error Correction (FEC)