

EMC.NCP-AII.v2026-06-17.q131

試験コード：	NCP-AII
試験名称：	NVIDIA AI Infrastructure
認証ベンダー：	EMC
無料問題の数：	131
バージョン：	v2026-06-17
ページの閲覧量：	102
問題集の閲覧量：	1362

<https://www.jpnsiken.com/shiken/EMC.NCP-AII.v2026-06-17.q131.html>

質問: 1

BlueField-2 DPUを2つの25GbEポートでリンクアグリゲーション (ボンディング)するように設定しています。ボンディングインターフェイスを設定した後、トラフィックが両方のリンクに分散されていないことに気づきました。この問題の最も可能性の高い原因を2つ選択してください。

- A. ボンディングモードが「balance-alb」に設定されており、ARP監視間隔が長すぎます。
- B. DPUに接続されているスイッチは、DPUで設定されているボンディングモードをサポートしていません。
- C. DPU上のファイアウォールが、インターフェースの1つで送信トラフィックをブロックしています。
- D. MTUサイズは、ボンディングインターフェースと物理インターフェースで異なります。
- E. 物理インターフェースの速度とデュプレックス設定が同じではありません。

正解: ([正解を表示します](#))

ボンディングには、サーバー (DPU) とスイッチの両方で互換性のある設定が必要です。スイッチがボンディングモードをサポートしていない場合、トラフィックは正しく分散されません。また、ボンディングが正しく機能するためには、ボンディング内のすべてのインターフェースで速度とデュプレックス設定が一致している必要があります。MTUが一致しないとパケットの断片化やパフォーマンスの問題が発生する可能性があります。トラフィックの分散が完全に妨げられる可能性は低くなります。ALBモードとARP監視はロードバランシングに影響を与える可能性があります。ファイアウォールがトラフィックをブロックする可能性はありますが、DPUの設定とはあまり関係がありません。

質問: 2

systemdを実行しているシステムで「nouveau」ドライバをブラックリストに登録し、NVIDIAドライバとの競合を防ぐには、どの設定ファイルを変更すればよいですか？

- A. /etc/modprobe.d/blacklist-nouveau.conf
- B. /etc/modules
- C. /boot/grub/grub.cfg

D. /etc/systemd/system.conf

E. /etc/default/grub

正解: ([正解を表示します](#))

ファイル `/etc/modprobe.d/blacklist-nouveau.conf` は、`nouveau` ドライバをブラックリストに登録する標準的な場所です。これにより、ロードされて NVIDIA ドライバと競合するのを防ぎます。`/etc/modules` は起動時にロードするモジュールのリストですが、ブラックリストには使用しません。`/boot/grub/grub.cfg` は GRUB 設定ファイルであり、モジュールのブラックリストには使用しません。`/etc/systemd/system.conf` は `systemd` 設定ファイルであり、モジュールとは関係ありません。`/etc/default/grub` は Grub 設定を構成して GRUB を更新する場所ですが、モジュールをブラックリストに登録する適切な場所ではありません。

質問: 3

Which of the following are crucial considerations when validating the hardware operation of an NVIDIA-Certified Professional AI Infrastructure server before deploying a production A1 workload? (Select all that apply)

- A. Ensuring the BIOS and BMC firmware are up to date.
- B. Verifying that the server's power supply units (PSUs) are operating within their specified voltage and current ranges under full load.
- C. Confirming the correct installation and functionality of all network interfaces (e.g., Ethernet, InfiniBand) relevant to the workload.
- D. Only testing the GPUs with a synthetic benchmark like 'linpack'.
- E. Checking if the server's case fans are spinning.

正解: ([正解を表示します](#))

Validating hardware operation involves checking BIOS/BMC firmware for compatibility and security updates, ensuring PSUs are within operating parameters, and verifying network interface functionality. Synthetic benchmarks alone aren't sufficient; the fans spinning is a basic check, but doesn't validate actual cooling performance under load. Complete hardware validation involves a broad range of checks to prevent later issues.

質問: 4

顧客から、既存のAIサーバーにNVIDIA GPUを追加インストールした後、GPUパススルーを利用するVMの起動時にIOMMU関連のエラーが発生しているとの報告がありました。エラーメッセージはアドレス変換の問題を示しています。どのトラブルシューティング手順を実行すべきでしょうか？ (2つ選択してください)

- A. BIOS/UEFIでIOMMIOを無効にしてください。これにより、アドレス変換の競合が解消されます。
- B. BIOS/UEFIでIOMMIOが有効になっていること、およびカーネルコマンドラインに `intel` CPUの場合は)または `amd` CPUの場合は)が含まれていることを確認してください。

- C. マザーボードのBIOS/UEFIを最新バージョンにアップデートしてください。新しいバージョンには、IOMMUのサポートの改善やバグ修正が含まれていることがよくあります。
- D. アドレス指定可能なメモリ制限を超えないように、VMに割り当てられるRAMの量を減らします。
- E. VM内でNVIDIAドライバーを再インストールします。

正解: ([正解を表示します](#))

IOMMUエラーは、多くの場合、設定ミスやファームウェアの古さが原因です。セキュリティが低下するため、IOMMUを無効にすることは一般的に推奨されません。BIOS/UEFIおよびカーネルコマンドラインでIOMMUが有効になっていることを確認することが重要です。BIOS/UEFIを更新することで、IOMMU実装における潜在的なバグに対処できます。メモリ割り当ても要因の一つとなり得ますが、最も可能性の高い直接的な原因ではありません。VM内でドライバーを再インストールしても、IOMMU関連のエラーが解決する可能性は低いでしょう。

質問: 5

AIインフラストラクチャは、複数のNVIDIA GPUから発生する熱を放散するために液体冷却システムに依存しています。最近のソフトウェアアップデート後、ユーザーから断続的なパフォーマンス低下とシステムクラッシュが報告されています。冷却の問題が疑われます。根本原因を診断する上で最も重要なチェック項目は次のうちどれですか？ (2つ選択)

- A. 液冷システム内のポンプ速度と冷却液流量を確認します。
- B. `sensors` コマンドを使用してCPU温度を確認します。
- C. システムログを分析して、GPU関連のエラー、特にサーマルスロットリングや電力制限を示すエラーを探します。
- D. データセンター内の周囲温度を調べます。
- E. ホストシステムでメモリテストを実行します。

正解: ([正解を表示します](#))

液冷システムにおいては、ポンプの速度と流量 (A)を確認することが非常に重要です。流量が不足すると、冷却不足や熱問題が発生する可能性があります。システムログを分析してGPU関連のエラー (C)を特定することで、冷却問題の一般的な症状であるサーマルスロットリングや電力制限が発生しているかどうかを直接確認できます。

質問: 6

AI/HPC環境でMIG (マルチインスタンスGPU)を使用する利点を正確に説明しているのは、次のうちどれですか？

(該当するものをすべて選択してください)

- A. MIGは、1つのGPUを複数の独立したインスタンスに分割することで、リソース利用率を向上させます。
- B. MIGは、同じGPU上で実行される異なるワークロード間のパフォーマンス分離を保証します。

- C. MIGはGPU仮想化ソフトウェアの必要性を排除します。
- D. MIGはGPUの全体的なFLOPSを向上させます。
- E. MIGでは、同じ物理GPU上で複数のCUDAバージョンを同時に実行できます。

正解: [\(正解を表示します\)](#)

MIGは、単一のGPUを複数の独立したインスタンスに分割することで、リソース利用率を最適化します (A)。ハードウェアレベルの分離を提供することで、パフォーマンスの一貫性を確保し、ワークロード間の干渉を防ぎます (B)。MIGはGPUの共有を簡素化しますが、特に複雑な環境では、すべての仮想化ソフトウェアの必要性を排除するものではありません。MIGはFLOPSを増加させるものではなく (D)、既存の計算能力を分割するものです。異なるMIGインスタンス上で実行される異なるコンテナは異なるCUDAバージョンを持つ可能性があります、MIG自体はこれを直接処理しません (E)。これはコンテナまたは環境を介して処理されます。

質問: 7

大規模なA1モデルを、ネットワーク接続ストレージ (NAS) デバイスに保存されたデータセットを使用してトレーニングしています。データ転送速度が予想よりも大幅に低下しています。初期トラブルシューティングの後、トレーニングサーバーとNASデバイスのネットワークインターフェイスのMTU (最大伝送単位サイズが一致していないことが判明しました。サーバーはMTUが1500に設定されていますが、NASデバイスはMTUが9000 (ジャンボフレーム) に設定されています。このMTUの不一致によって最も起こりうる結果は何ですか？また、どのような対策を講じるべきですか？

- A. データパケットが断片化され、オーバーヘッドが増加し、パフォーマンスが低下します。サーバーとNASデバイスの両方で、同じMTUサイズ (1500または9000) を使用するように設定してください。
- B. サーバーとNASデバイス間の接続が不安定になり、データ破損が発生する可能性があります。両方のデバイスのMTUサイズを、サポートされている最大値まで増やしてください。
- C. サーバーがNASデバイスと通信できなくなります。サーバーのMTUサイズをNASデバイスのMTUサイズに合わせて小さくしてください。
- D. データ転送は最小共通MTUサイズに制限されますが、パフォーマンスに大きな影響はありません。特に対応は必要ありません。
- E. データパケットが再送信されるため、遅延は増加しますが、スループットは維持されません。サーバーをパスMTU検出 (PMTUD) を使用するように設定してください。

正解: [\(正解を表示します\)](#)

An MTU mismatch (option A) will cause fragmentation, where larger packets are broken down into smaller packets before being transmitted, adding overhead and reducing performance. The solution is to configure both devices to use the same MTU size. Choosing 1500 ensures compatibility, while 9000 requires the entire network path to support jumbo frames.

質問: 8

NVIDIA AIOO GPU 上で MIG を構成した後、`nvidia-smu`を実行すると、すべての MIG インスタンスが「無効」状態になっていることがわかります。

この問題の考えられる原因として、以下のうちどれが挙げられますか？（該当するものをすべて選択してください）

- A. 必要なNVIDIAドライバーがインストールされていないか、GPUと互換性がありません。
- B. GPUがMIGモードではありません。
- C. MIGインスタンスは正しく構成されていますが、どのプロセスにも割り当てられていません。
- D. 'nvidia-persistenced' サービスが実行されていません。
- E. システムの電源供給が不足しています。

正解: A,B,D ([コメントを发表する](#))

MIGインスタンスが「無効」状態になっている場合、MIG構成に根本的な問題があることを示しています。互換性のないドライバ A)は、インスタンスの適切な初期化を妨げません。GPUが明示的にMIGモードに設定されていない場合 B)、MIGインスタンスは利用できません。`nvidia-persistenced`サービス D)は、ドライバの状態が再起動後も維持されるようにしますが、このサービスがないとMIGインスタンスが無効状態に戻る可能性があります。割り当てられていないインスタンス C)は存在しますが、「無効」状態ではなく「アイドル」状態であるべきです。電源供給が不十分な場合 E)はGPUが正しく動作しない可能性があります。MIGが「無効」状態になる可能性は低いでしょう。

質問: 9

最近のCUDAアップデート後、マルチGPU通信にNCCL (NVIDIA Collective Communications Library)を使用するディープラーニングのトレーニングジョブが「NCCL 内部エラー、未処理のシステムエラー」で失敗するようになりました。このエラーは `all reduce`操作中に発生します。

最も可能性の高い根本原因は何ですか？また、どのように対処しますか？

- A. NCCLのバージョンが新しいCUDAバージョンと互換性がありません。インストールされているCUDAバージョンと互換性のあるNCCLバージョンにアップデートしてください。
- B. CUDA コンテキストに割り当てられた共有メモリが不足しています。`'cudaDeviceSetLimit(cudaLimitSharedMemory, new_limity)'` を使用して共有メモリ制限を増やしてください。
- C. ファイアウォールルールがGPU間の通信をブロックしています。ファイアウォールを設定して、NCCLで定義されたポート（通常は8000~8010）での通信を許可してください。
- D. ノード間通信に使用されているネットワークケーブルに不具合がある（トレーニングジョブが複数のサーバーにまたがる場合）。ネットワークケーブルを認証済みの高速ケーブルに交換してください。
- E. GPU Direct RDMAが正しく設定されていません。`ibmsg`でエラーを確認し、RDMAが有効になっていることを確認してください。

正解: **A** ([コメントを发表する](#))

NCCLは特定のCUDAバージョンに依存しています。CUDAアップデート後の互換性の問題が最も可能性の高い原因です。共有メモリ不足がNCCL内でシステムエラーを引き起こす可能性は低いです。ファイアウォールルールは通常、接続拒否エラーとして現れます。ネットワークケーブルの不具合はノード間の通信に影響を与え、ノード内の通信には影響しません。RDMAの問題は問題を引き起こす可能性があります、通常はCUDAアップデート直後に「未処理のシステムエラー」として現れることはなく、RDMAが以前正常に動作していた場合に発生する可能性が高くなります。

質問: 10

8基のGPUを搭載したAIサーバーで、高負荷時にランダムなシステムクラッシュが発生しています。システムログにはメモリエラーの可能性が示されていますが、標準的なメモリテスト (memtest86+)はエラーなく合格しています。GPUはパッシブ冷却方式です。これらのクラッシュの最も可能性の高い根本原因を3つ挙げてください。

- A. インストールされているLinuxカーネルと互換性のないNVIDIAドライバーバージョンです。
- B. 標準的なCPUベースのメモリテストでは検出できないGPUメモリエラー。
- C. サーバー内部の空気の流れが不十分なため、GPUとVRMが過熱します。
- D. ピーク負荷時に安定した電力を供給できない電源ユニット (PSU)の故障。
- E. 分散トレーニング中にネットワークの輻輳により断続的なデータ破損が発生します。

正解: **B,C,D** ([コメントを发表する](#))

CPUベースのテストではGPUメモリを直接テストしないため、GPUメモリのエラー (B)が原因の可能性が高いです。受動冷却方式のため、エアフローが不十分 (C)になっている可能性が高く、熱不安定を引き起こしています。電源ユニット (PSU)の不具合 (D)は、負荷がかかった際に電力変動によってランダムなクラッシュを引き起こす可能性があります。ドライバの互換性の問題 (A)は、初期設定後にランダムなクラッシュを引き起こす可能性は低く、ネットワークの混雑 (E)は通常、システムクラッシュではなくトレーニングの速度低下につながります。

質問: 11

NVIDIA GPUを搭載したサーバーを最適な電力効率で構成しています。アイドル時のエネルギー消費を最小限に抑えるために、NVIDIAの電力管理機能を活用したいと考えています。アクティブなワークロード中のパフォーマンスに大きな影響を与えることなく、この目標を達成するために最も効果的なアクションは次のうちどれでしょうか？

- A. ワークロードに関係なく、GPUのクロック速度を可能な限り低い設定に下げます。
- B. NVIDIAのアダプティブクロッキングと電力制限機能を有効にして、ワークロードに基づいてGPUがクロック速度と消費電力を動的に調整できるようにします。
- C. 常に最高のパフォーマンスを確保するために、すべてのGPU電源管理機能を無効にします。

- D. サーバーから1つ以上のGPUを取り外して、全体の消費電力を削減します。
- E. GPUの静的電力制限を非常に低く設定し、アクティブなワークロード中でもパフォーマンスを大幅に制限します。

正解: [\(正解を表示します\)](#)

NVIDIAのアダプティブクロッキングと電力制限機能を有効にするのが最も効果的な方法です。これらの機能により、GPUはワークロードに基づいてクロック速度と消費電力を動的に調整し、アイドル時のエネルギー消費を最小限に抑えつつ、アクティブなワークロード時のパフォーマンスを最大限に高めます。クロック速度を低く固定する (A)か、電力制限を設定する (E)と、パフォーマンスに深刻な影響が出ます。電源管理を無効にする (C)と、エネルギーが無駄になります。GPUを取り外す (D)と、パフォーマンスが低下します。

質問: 12

A1インフラストラクチャでNVIDIA Spectrum-Xスイッチを使用しています。大規模な分散トレーニングジョブ中に、2台のGPUサーバー間で高いレイテンシが発生していることがわかりました。スイッチのテレメトリを分析した結果、ルーティングパスが最適ではないことが問題の原因となっている可能性があるかと推測されます。この問題を軽減するために、Spectrum-Xファブリック内のトラフィックフローを最も細かく制御できる方法はどれですか？

- A. すべてのスイッチで、等コストマルチパス (ECMP) ハッシュアルゴリズムをグローバルに調整します。
- B. QoS (Quality of Service) ポリシーを設定して、レイテンシの高いGPUサーバーからのトラフィックを優先します。
- C. Spectrum-Xスイッチで利用可能なアダプティブルーティング (AR) またはダイナミックロードバランシング (DLB) 機能を実装し、ネットワークの状態に基づいてパスを動的に調整します。
- D. Spectrum-X スwitchで静的ルートを手動で設定し、GPUサーバー間のトラフィックを特定のパスに強制します。
- E. ルーティングの決定を簡素化するために、IPv6を無効にします。

正解: [\(正解を表示します\)](#)

アダプティブルーティング (AR) とダイナミックロードバランシング (DLB) は、Spectrum-Xのリアルタイムネットワーク状況に基づいてパスを動的に調整するために特別に設計された機能です。これにより、静的ルーティングやグローバルECMP調整と比較して、輻輳への対応とトラフィックフローの最適化を最もきめ細かく自動化された方法で実現できます。QoSは優先順位付けを行います。選択されたパス自体は変更しません。

質問: 13

複数のラックにまたがる大規模なAIトレーニングクラスターを設計しています。ネットワークポロジでは、ラック内 (短距離) 接続ラック間 (長距離) 接続両方が必要で

す。200GbEと400GbEリンクが混在する場合、このシナリオにおいて最もコスト効率が高く、適切なケーブルの種類とトランシーバーの組み合わせはどれでしょうか？

- A. ラック内接続にはDACケーブル、ラック間接続にはAOCケーブルを使用します。
- B. ラック内接続にはDACケーブル、ラック間接続にはマルチモードファイバーとシングルモードファイバーに対応したSR4とLR4のトランシーバーをそれぞれ組み合わせて使用します。
- C. ラック内およびラック間のすべての接続に対応するAOCケーブル。
- D. ラック内接続にはすべてパッシブ銅ケーブル、ラック間接続にはシングルモードファイバー付きER4トランシーバーを使用。
- E. ラック内接続用のDACケーブルと、ラック間接続用のシングルモードファイバー付きFR4トランシーバー。

正解: [\(正解を表示します\)](#)

DACケーブルはコスト効率が高く、ラック内の短距離 高帯域幅接続に適しています。ラック間接続には、マルチモードファイバーを使用したSR4トランシーバー（短距離ラック間リンク用）と、シングルモードファイバーを使用したLR4トランシーバー（長距離ラック間リンク用）が、コストと性能のバランスに優れています。AOCは一般的にDACよりも高価です。ER4は多くのラック間接続シナリオにおいてオーバースペックであり、LR4よりも高価です。

質問: 14

InfiniBandホストで使用される次の `ibroute` コマンドを考えてみましょう。 `ibroute add dest Ox1a dev ib0`。このコマンドの最も可能性の高い目的は何ですか？

- A. InfiniBandサブネット外宛てのすべてのトラフィックに対してデフォルトルートを追加します。
- B. InfiniBandインターフェイス `ib0` を使用して、LID `Ox1a`宛のトラフィックの静的ルートを作成します。
- C. `ib0` インターフェースのMTUサイズを `0x1a` バイトに設定します。
- D. `ib0` インターフェースでのルーティングを無効にします。
- E. InfiniBandインターフェイス `ib0` を使用して、IPアドレス `Ox1a`宛のトラフィックの静的ルートを設定します。

正解: **B** ([コメントを发表する](#))

`ibroute add dest Ox1a dev ib0` コマンドは、InfiniBand LID（ローカル識別子）`Ox1a`宛てのトラフィックに対して、InfiniBandインターフェイス `ib0` を使用した静的ルートを作成します。InfiniBandルーティングは主にLIDに基づいており、IPアドレスを直接使用するわけではありません（ただし、IP over 1Bも可能です）。 `dest` パラメータは宛先LIDを指定します。

質問: 15

ノードごとに複数のNVIDIA AIOO GPUを搭載した新しいクラスターをデプロイします。NVLinkを使用してGPU間の通信パフォーマンスを最適化したいと考えています。NVLinkの帯域幅を最大化するために、以下の構成のうちどれが重要ですか？

- A. ノード内のすべてのGPUは、同じモデルで、ファームウェアのバージョンも同一である必要があります。
- B. NVLinkの帯域幅を最大化するには、マザーボードがPCIe Gen5をサポートしている必要があります。
- C. GPUは、サーバーのアーキテクチャに基づいて、NVLinkの直接接続を最大化するスロットに物理的にインストールする必要があります。
- D. NVIDIA ドライバーで NVLink を有効にする必要があります。デフォルトでは無効になっています。
- E. NVLink機能を利用するには、サーバーは特定のCPUモデルを使用する必要があります。

正解: [\(正解を表示します\)](#)

NVLink のパフォーマンスを最適化するには、いくつかの条件を満たす必要があります。同じモデルとファームウェアの GPU を使用することで互換性が確保され、パフォーマンスのボトルネックを防ぐことができます。物理的な配置も重要です。サーバーのアーキテクチャとドキュメントで定義されているように、NVLink の直接接続を最大化するスロットに GPU をインストールする必要があります。PCIe Gen5 はシステム全体のパフォーマンス向上に役立ちますが、NVLink の帯域幅に直接影響を与えるわけではありません。NVLink は通常、デフォルトで有効になっています。一部の CPU モデルは推奨される場合もありますが、より重要なのはマザーボードの NVLink トポロジです。

質問: 16

ユーザーが複数の組織にまたがるリソースにアクセスできるようにNGC CLIを設定する必要があります。これを実現するための推奨コマンド構文は何ですか？

- A. `export NGC_CLI_ORG=org-name && ngc config set`
- B. `ngc config list` を使用して JSON 設定ファイルを手動で編集します。
- C. `ngc registry login --org org-name`
- D. `ngc config set --org org-name --ace ace-name`

正解: [D \(コメントを发表する\)](#)

正解は `ngc config set --org org-name --ace ace-name` です。NGC CLI は構成コンテキストを使用して、ユーザーが操作している組織、チーム、または権限スコープを判断します。ユーザーが複数の組織に所属している場合、レジストリ アクセス、モデル ダウンロード、プライベート コンテナ、およびエンタープライズ ソフトウェア リソースが正しく解決されるように、CLI を目的の組織コンテキストで構成する必要があります。NVIDIA のドキュメントでは、`ngc config set` が組織とチームの値を含む NGC コンテキストを構成または切り替えるために使用されるコマンドとして示されており、他の NVIDIA インストール ワークフローでは、NGC 構成で組織、チーム、および ACE フィールドの入力が求められます。JSON 構成ファイルを手動で編集することは、構文エラーや認証状態の不整合のリスク

が高まるため、推奨される操作方法ではありません。ngc registry login は、永続的な NGC CLI 組織コンテキストを設定する通常のコマンドではありません。環境変数をエクスポートするとセッションに影響する可能性があります。永続的な構成方法としては最適ではありません。AI インフラストラクチャの操作では、検証済みのコンテナ、GPU ドライバー、vGPU ソフトウェア、モデル、およびエンタープライズ アセットを正しいアカウントスコープから取得するために、正しい NGC 構成が重要です。

有効的な**NCP-AII**問題集はJPNTTest.com提供され、**NCP-AII**試験に合格することに役に立ちます！JPNTTest.comは今最新**NCP-AII**試験問題集を提供します。JPNTTest.com NCP-AII試験問題集はもう更新されました。ここで**NCP-AII**問題集のテストエンジンを手に入れます。最新版のアクセス、<https://www.jpntest.com/shiken/NCP-AII-mondaishu> **125問、30%ディスカウント**、特別な割引コード: **JPNshiken**」

質問: 17

NVLinkで接続された2つのNVIDIA GPUを搭載したサーバーがあります。NVLinkが正しく機能していることを確認したい場合、NVLinkの状態と帯域幅を確認するために使用できるコマンドまたはツールは何ですか？

- A. 'nvidia-smi nvlink -statue
- B. 種
- C. nvcc -version'
- D. 'nvidia-smi topo -m'
- E. 'nvidia-settings' (GUIツール)

正解: ([正解を表示します](#))

'nvidia-smi nvlink -status' は、リンク速度やエラーなど、NVLink の状態を直接概観できません。'nvidia-smi topo' は、NVLink 接続を含む GPU のトポロジと接続方法を表示します。'lspci' は PCL デバイスを一覧表示しますが、NVLink 固有の情報は提供しません。'nvcc -version' は CUDA コンパイラのバージョンを確認します。'nvidia-settings' は、いくつかの情報を表示できる GUI ツールですが、NVLink の状態については 'nvidia-smr' ほど正確ではありません。

質問: 18

企業プロキシ経由で pip]を使用してNGC CLIをインストールしようとしていますが、接続エラーのためインストールが失敗します。NGC CLIのインストール中にプロキシを使用するように pip]を設定するにはどうすればよいですか？

- A. プロキシサーバーのアドレスとポートを使用して、http_proxy]と https_proxy]環境変数を設定します。
- B. 'pip install' コマンドで '-proxy' オプションを使用し、プロキシサーバーのアドレスとポートを指定します。

- C. 適切なディレクトリ (例: `~/pip/pip.conf`) に `pip.conf` ファイルを作成し、ファイル内でプロキシ設定を構成します。
- D. NGC CLIパッケージを手動でダウンロードし、オフラインでインストールします。
- E. NGC CLIはプロキシの使用をサポートしていません

正解: [\(正解を表示します\)](#)

環境変数を設定する A)、`pip install` コマンドで `-proxy` オプションを使用する B)、または `pip.conf` ファイルを作成する C) ことで、`pip` をプロキシを使用するように構成できます。パッケージを手動でダウンロードする D) 方法は、初期インストール時の回避策ですが、NGC との継続的な通信には対応していません。NGC CLI はプロキシ構成をサポートしているため、オプション E は誤りです。

質問: 19

256個のGPUを搭載したNVIDIAエンタープライズAIファクトリーにおいて、スケーリングテスト中に検証すべき最も重要なストレージソリューションの特性は何ですか？

- A. ノードごとのスループットが8 GiB/s以上で安定していること。
- B. アイドル状態のクラスタにおけるシングルノード書き込みパフォーマンス。
- C. ディスク障害時のRAID再構築時間。
- D. 最大4Kランダム読み取りIOPSが100万を超えます。

正解: [\(正解を表示します\)](#)

AIクラスタを256個のGPU (DGX H100の32ノード)に拡張すると、ストレージファブリックに大規模な「インキャスト」問題が発生します。大規模なトレーニング中、各ノードは頻繁に大量のデータを同時に読み込みます。

NVIDIA のリファレンス アーキテクチャ (BasePOD/SuperPOD) では、高性能トレーニングの場合、各ノードが8個のGPUすべてを飽和状態に保つために、最低限のスループット(多くの場合8 GiB/s以上)を維持できる必要があると規定されています。ストレージシステムが1つのノードを高速で処理できるものの、32個のノードすべてがデータを要求すると処理が滞ると、GPUがIOを待ってアイドル状態になるため、AIモデルの「スケーリング効率」が大幅に低下します。したがって、フルクラスタ負荷下でのノードごとの一貫したスループットを検証することが、AIファクトリーにとって最も重要な指標となります。IOPS(オプションD)は小さなファイルでは重要ですが、最新のAIデータセットは多くの場合、WebDatasetやTFRecordのような大きなバイナリフォーマットにシャーディングされるため、シーケンシャルスループットが主なボトルネックとなります。

質問: 20

ユーザーから、ディープラーニングのトレーニングジョブが「CUDAメモリ不足」エラーでクラッシュするという報告がありました。ただし、`nvidia-smi` コマンドではGPUに十分な空きメモリがあると表示されています。このジョブはTensorFlowを使用しています。考えられる原因は2つあります。

- A. TensorFlow のバージョンは、インストールされている NVIDIA ドライバーと互換性がありません。
 - B. TensorFlow は GPU ではなく CPU にメモリを割り当てています。
 - C. TensorFlow が GPU メモリを断片化しているため、連続したブロックを割り当てるのが困難になっています。
 - D. CUDA_VISIBLE_DEVICES 環境変数が正しく設定されていません。
 - E. システムのスワップ領域がいっぱいなので、メモリを割り当てることができません。
- 正解: ([正解を表示します](#))

一見GPUメモリが利用可能に見えても、CUDAメモリ不足エラーが発生する場合は、メモリの断片化またはGPUの割り当てが不適切であることを示しています。TensorFlowはGPUメモリを断片化することがあり、合計メモリが十分にある場合でも割り当てエラーが発生する可能性があります。この変数は、TensorFlowがアクセスできるGPUを制御します。この変数が設定されていないか、誤って設定されている場合、TensorFlowは存在しない、または利用できないGPUにメモリを割り当てようとしている可能性があります。TensorFlowのバージョン非互換性は問題を引き起こす可能性があります、直接「CUDAメモリ不足」エラーとして現れる可能性は低くなります。TensorFlowは、正しく構成されていれば、通常GPUメモリの割り当てを優先します。

質問: 21

サーバーにNVIDIA AIO GPUを構成していますが、インストールとドライバの設定後、GPUの指定TDPよりも低い値になっています。考えられる原因は何ですか？nvidia-smiは電力制限を報告しています。

- A. GPUに不具合があります。
- B. 電源装置が十分な電力を供給していません。
- C. ドライバーが正しくインストールされていません。
- D. システムBIOSがPCIeスロットへの電力供給を制限しています。
- E. GPUは非アクティブ状態のため低電力モードになっています。

正解: **D** ([コメントを发表する](#))

他の選択肢も考えられますが、BIOS設定でPCIeスロットへの電力供給が制限されていることが、nvidia-smiで報告される電力制限値が予想外に低い場合の一般的な原因です。電力関連の問題のトラブルシューティングを行う際は、必ずBIOS設定を確認してください。GPUが低電力モードの場合、負荷がかかると電力供給が増加するはずですが。

質問: 22

新しいGPUクラスタでNVIDIA Base Command Manager (BCM)の高可用性(HA)をセットアップするタスクが割り当てられました。クラスタは、プライマリヘッドノード、セカンダリヘッドノード、および複数の計算ノードで構成されています。要件は、BCMサービスの自動フェイルオーバー、ワークロードへの影響の最小化、およびインストール中とインストール後のクラスタの健全性監視の適切な実施です。BCM HAのインストールと構成プロ

セスにおいて、堅牢で検証済みのHAクラスタ構成を確保するために必須となるアクションは次のうちどれですか？ (2つ選択)

以下の選択肢から正しいものを2つ選んでください。

- A. フェイルオーバー時にプライマリヘッドノードとセカンダリヘッドノード間で自動的に移行できるフローティング仮想IPアドレスを割り当てます。
- B. ヘッドノードの同期を開始するには、計算ノードの電源がオンになっていて、作業を実行している必要があります。
- C. 設定が完了したら、アクティブなヘッドノードで BCM サービスを停止してフェイルオーバーをシミュレートし、セカンダリノードですべてのサービスが中断なく実行されていることを確認します。
- D. 両方のヘッドノードで、共有仮想IPアドレスに依存するのではなく、BCMサービスに独立した静的IPアドレスを使用するように設定します。
- E. 設定時に、プライマリヘッドノードからセカンダリヘッドノードへ設定データディレクトリと状態データディレクトリの両方を明示的に同期して、一貫性を確保します。

正解: A,C ([コメントを发表する](#))

ユーザー、コンピューティング ノード、および管理サービスがアクティブな BCM ヘッドノードに対して一貫したエンドポイントを必要とするため、フローティング仮想 IP アドレスが必要です。NVIDIA DGX BasePOD HA のドキュメントには、外部 VIP は HA 構成が完了した後にアクティブなヘッドノードにアクセスするために常に使用する必要がある IP アドレスであると記載されています。同じ HA ワークフローでは、セットアップ後にフェイルオーバー動作も検証し、双方向から HA ステータスを確認し、どのヘッドノードがアクティブであるかを確認します。したがって、セカンダリヘッドノードがサービス責任を引き継ぐことができることを証明するには、制御されたフェイルオーバー検証が必要です。HA セットアップ中は、コンピューティング ノードの電源をオンにしてワークロードを実行する必要はありません。NVIDIA のドキュメントでは、HA 構成の前にクラスタ ノードの電源をオフにする例が示されています。独立した静的 IP アドレスだけでは、クライアントが手動で再接続する必要があるため、透過的なサービス継続性は提供されません。同期はHAセットアップワークフローの一部として行われますが、自動フェイルオーバーと検証済みの準備状況を実現するための最適な2つの方法は、フローティングVIPとフェイルオーバー検証です。

質問: 23

あなたは、AIワークロードに大きく依存するデータセンターのエネルギー効率を向上させるプロジェクトを率いています。NVIDIAは、最新のデータセンターの効率をより適切に把握するために、電力使用効率 (PUE) などの従来の指標を超えたアプローチを提案しています。どちらの戦略を優先すべきでしょうか？

- A. 電力使用効率を主要な指標として使用し、エネルギー単位当たりの有効仕事量などの追加指標でそれを補完する。

- B.** ワット数を効率の主要な指標として使用します。これは、任意の時点での電力入力を正確に反映するためです。
- C.** AIアプリケーション向けのMLPerfなど、特定のワークロードに合わせたベンチマークを開発し、実際のシナリオにおけるエネルギー使用量をよりよく理解する。
- D.** 生産的な作業に実際に使用されたエネルギーをより正確に反映するために、キロワット時を既存の指標に統合することに重点を置く。

正解: [\(正解を表示します\)](#)

Traditional data center metrics like PUE (Power Usage Effectiveness) only measure how much energy is "wasted" by cooling and power delivery relative to the IT load; they say nothing about how efficiently that IT load is performing its task. In an AI Factory, "Efficiency" is better defined by the amount of AI training or inference performed per watt. NVIDIA advocates for the use of workload-specific benchmarks, such as MLPerf, to quantify this. MLPerf measures the time and energy required to complete standardized AI tasks (like training a ResNet-50 model or an LLM). By prioritizing these benchmarks (Option C), an organization can compare the energy efficiency of different hardware architectures (e.g., A100 vs. H100) or different software optimizations (e.g., FP8 vs. FP16). For example, even if an H100 system draws more peak power than an older system, its ability to complete a training job 9x faster results in a significantly lower "Total Energy Consumed per Job". This shift from "infrastructure efficiency" (PUE) to "computing efficiency" (MLPerf-per-watt) is essential for modern AI data centers aiming for sustainability and cost-effective scaling.

質問: **24**

You are deploying a BlueField-3 DPU within a secure environment. You are required to enable secure boot to prevent unauthorized firmware from being loaded. Which steps are typically involved in enabling secure boot on a BlueField-3 DPU, starting from a factory-default state? (Select TWO)

- A.** Install the latest version of the Mellanox OFED drivers on the host server.
- B.** Generate cryptographic keys and enroll them in the DPIJ's IJEFI firmware using tools provided by NVIDIA or Mellanox.
- C.** Configure the server's BIOS to enable UEFI boot mode.
- D.** Flash the BlueField-3 DPIJ with a secure boot-enabled OS image that is signed with the enrolled keys.
- E.** Update the DPU's BMC firmware to the latest version.

正解: **B,D** ([コメントを发表する](#))

重要な手順は、信頼チェーンを確立する暗号鍵を生成して DPIJ の UEFI ファームウェアに登録すること、そしてそれらの鍵で署名されたセキュアブート対応の OS イメージを DPIJ に書き込むことです。これらがなければ、セキュアブートは正しく機能しません。BMC ファームウェア (E) のアップデートは一般的に推奨される方法であり、サーバーの BIOS (C) で IJEFI ブートモードを有効にすることは UEFI 全般に必要ですが、これらは DPIJ 自

体でセキュアブートを有効にするための具体的な手順ではありません。ホストドライバのインストール (A) もセキュアブートとは無関係です。

質問: 25

顧客はエンタープライズ規模の展開を想定したAIファクトリーを設計しており、管理ネットワークとストレージネットワークの冗長性と負荷分散を確保したいと考えています。イーサネットスイッチにはどの機能を実装すべきでしょうか？

- A. スパニングツリープロトコルを使用して冗長スイッチを実装します。
- B. 冗長スイッチ間でボンディングされたインターフェースのためのMLAG。
- C. 管理およびストレージトラフィックすべてに1つのスイッチのみを使用します。
- D. VLANを無効にして、管理機能のないスイッチを使用します。

正解: ([正解を表示します](#))

NVIDIA AIファクトリーの「南北」および「管理/ストレージ」イーサネットファブリックでは、高可用性が最重要事項です。独自のルーティングロジックを使用するInfiniBandコンピューティングファブリックとは異なり、イーサネット側は標準的なデータセンタープロトコルに依存しています。真のハードウェア冗長性を提供し、利用可能な帯域幅を2倍にする（負荷分散ために、NVIDIAはMLAG（マルチシャーシリンクアグリゲーション）を推奨しています。

MLAGを使用すると、2台の物理スイッチをDGXノードから単一の論理ユニットとして認識できます。DGXは、2つのイーサネットNICをボンディング（例えば、802.3ad LACPボンディング）し、各スイッチに1本のケーブルを接続できます。この構成にはいくつかの利点があります。1台のスイッチが故障した場合、トラフィックはスパニングツリープロトコル（オプションA）に伴う遅い収束時間なしに、シームレスに他のリンクにとどまります。さらに、クラスタは、NFSやS3の取り込みなどの大容量ストレージトラフィックに対して、両方のリンクの帯域幅をまとめて利用できます。単一のスイッチ（オプションC）または管理されていないハードウェア（オプションD）を使用すると、単一障害点が発生し、セキュアなAIインフラストラクチャに必要なトラフィック分離（VLAN）が欠如します。

質問: 26

企業のITチームは、すべてのGPUサーバーにSpectrum-X Ethernetネットワークを接続したAIファクトリーの物理的な設置を完了しました。次に、スケーラブルなAIワークロードの展開に対応できる環境が整っていることを確認する必要があります。推奨される検証手順の順序は何ですか？

- A. Active DirectoryとLDAPを設定し、ロールベースのアクセス制御とセキュリティ設定を最初に構成し、ユーザーをインストールし、ネットワークまたはハードウェアのパフォーマンス検証をスキップします。
- B. まずアプリケーションのベンチマークを実行し、パフォーマンスログを使用してボトルネックを特定し、その後スイッチとサーバーのファームウェアを更新し、パフォーマンステストを使用してネットワークを調整します。

C. ソフトウェアスタックを検証し、リンク接続とポートの状態をテストし、ネットワークベンチマークを実行し、OSPFを実行し、ネイバーがルート情報を交換していることを確認し、AIワークロードテストを実施します。

D. スイッチとサーバーのファームウェア構成を確認し、リンク接続とポートの状態をテストし、ネットワークベンチマークを実行し、ソフトウェアスタックを検証してから、AIワークロードテストを実施します。

正解: [D \(コメントを发表する\)](#)

正しい手順は、まずスイッチとサーバーのファームウェア構成を確認し、次にリンク接続とポートの状態を検証し、ネットワークベンチマークを実行し、ソフトウェアスタックを検証し、最後にAIワークロードテストをステージングすることです。この順序は、階層化されたインフラストラクチャ検証アプローチに従います。ファームウェアと構成は最初に検証する必要があります。ファームウェアの不一致やサポートされていない設定は、リンクの不安定性、RoCE動作の誤り、または輻輳制御パフォーマンスの一貫性の欠如を引き起こす可能性があるためです。次に、リンクとポートのチェックを実行して、すべての物理的および論理的な接続が動作していることを確認します。次に、ネットワークベンチマークを実行して、スループット、レイテンシ、パケット動作、およびGPU間通信の準備状況を検証します。ファブリックが健全であることが証明された後のみ、ソフトウェアスタックとAIワークロードのステージングを検証する必要があります。ファームウェアとリンクの検証の前にアプリケーションベンチマークを開始すると、障害が下位レイヤーの問題によって引き起こされる可能性があるため、トラブルシューティングが難しくなります。LDAPやRBACなどのIDサービスは運用には重要ですが、ファブリックの準備状況を証明するものではありません。Spectrum-X AI Ethernetの導入において、拡張可能なワークロードへの対応は、クリーンな物理リンク、適切なファームウェア、健全なポートテレメトリ、検証済みのネットワークパフォーマンス、そしてアプリケーションレベルのテストに依存します。

質問: 27

データセンターは、東西方向のトラフィックが多いAIトレーニング向けに設計されています。コストとパフォーマンスを考慮すると、一般的にどのネットワークトポロジーが最も適しているのでしょうか？

A. 背葉

B. 3段

C. リング

D. バス

E. メッシュ

正解: [\(正解を表示します\)](#)

スパインリーフアーキテクチャは、AIトレーニングに特徴的な高帯域幅 低遅延のトラフィックパターンに対応するように設計されています。任意の2つのサーバー間でコストが等しいパスを持つノンブロッキングファブリックを提供するため、東西方向の通信に最適です。3層アーキテクチャは、南北方向のトラフィックを伴う従来型のアプリケーションに

適しています。リング型とバス型は拡張性が低く、高負荷時にはパフォーマンスが低下します。メッシュ型は複雑で、大規模な展開にはコストがかかります。

質問: 28

AIクラスタは、シングルモード光ファイバーを使用して2kmの距離で200Gbpsのデータを送信する必要があります。コストとパフォーマンスを考慮すると、どのトランシーバタイプが最も適切でしょうか？

- A. 200GBASE-SR4
- B. 200GBASE-LR4
- C. 200GBASE-ER4
- D. 200GBASE-CR4
- E. 200GBASE-DR4

正解: ([正解を表示します](#))

200GBASE-LR4が最も適しています。LRはLong Reach（長距離を意味し、通常はシングルモードファイバーで最大10kmまでです。SRはShort Reach（短距離）（通常銅線または非常に短いファイバー）です。ERはExtended Reach（長距離）（最大10km）です。CRは銅線で、非常に短い距離用です。DRは通常、具体的な実装に応じて最大500mから2kmまでの距離で使用されるため、LRは可能ではありますが、確実に2kmの距離が必要な場合はより適しています。

質問: 29

NVIDIA Spectrum-X ベースの A1 クラスタでネットワーク パフォーマンスの問題をトラブルシューティングしています。ECMP (Equal-Cost Multi-Path) ハッシュ アルゴリズムが利用可能なパス全体にトラフィックを均等に分散しておらず、一部のリンクで輻輳が発生しているのではないかと疑っています。この問題を検証して対処するために、次のうちどの方法が最も効果的でしょうか？

- A. 影響を受けるノード間でパケットが通過する経路を分析するには、「ping」または「fraceroute」を使用してください。パケットが常に同じ経路を通る場合は、ECMPが正しく機能していない可能性があります。
- B. スイッチのテレメトリツール (NVIDIA What's Up Gold、Mellanox NEOなど) を使用して、ノード間の利用可能なすべてのパスにおけるリンク使用率を監視します。トラフィック量の著しい不均衡を探します。
- C. スイッチを再起動して、ECMPハッシュアルゴリズムにパスを再計算させます。
- D. ECMPを完全に無効にして、静的ルーティングのみに依存します。
- E. TCPウィンドウサイズを小さくします。

正解: ([正解を表示します](#))

スイッチのテレメトリツールは、リンク使用率を監視し、ECMPによって引き起こされるトラフィック分配の不均衡を特定するための最も直接的かつ包括的な方法を提供します。

「ping」や「fraceroute」はパス情報を提供できますが、トラフィック量に関する情報は得られ

ません。スイッチを再起動すると一時的に問題が緩和される可能性があります、ECMP ハッシュの根本的な問題は解決されません。ECMPを無効にするのは最終手段であり、全体の帯域幅が低下する可能性があります。

質問: 30

あるメディア企業は、大量の非構造化ビデオデータを保存 処理する必要のある、ビデオコンテンツ分析のためのAIプラットフォームを開発している。このプラットフォームは、データ取り込みのための高いスループットをサポートし、リアルタイム分析のための効率的なアクセスを提供する必要がある。これらの要件を踏まえ、同社はどのようなストレージ戦略を採用すべきだろうか？

- A. コスト効率とアーカイブ機能に優れたテープストレージ
- B. 低遅延と高性能を実現するブロックストレージ
- C. 階層構造による整理と容易なナビゲーションを実現するファイルストレージ
- D. スケーラビリティとメタデータ管理のためのオブジェクトストレージ

正解: ([正解を表示します](#))

オブジェクトストレージは大規模なデータやメタデータには最適ですが、NVIDIA AI インフラストラクチャのトレーニングワークロード、特にビデオ分析のベストプラクティスでは、並列ファイルシステム (PFS) を最優先しています。最新の AI フレームワーク (PyTorch、TensorFlow) や NVIDIA 独自の SDK (DeepStream や NeMo など) は、POSIX 準拠のファイルシステムから読み取るように最適化されています。ビデオコンテンツ分析の場合、トレーニングプロセスでは、大きなビデオファイルを「シャーディング」し、大規模なデータセットに対してランダムアクセス読み取りを実行します。高性能ファイルシステム (Lustre、Weka、IBM Storage Scale など) は、ノードあたり 8 台以上の H100 GPU をデータで飽和状態に保つために必要な、高スループットかつ低遅延のメタデータ操作を提供します。ファイルストレージは、データサイエンティストがデータセット (例: /datasets/train/videos/) を管理するために使用する階層構造を可能にし、GPU がストレージファブリックから GPU メモリに直接データをプルして CPU をバイパスし、取り込みスループットを最大化できる GPUDirect Storage (GDS) をサポートします。

質問: 31

InfiniBandファブリックにおいて、ルーティングに関してサブネットマネージャ (SM)の主な役割は何ですか？

- A. 従来のIPルーターと同様に、宛先IPアドレスに基づいてパケットを転送します。
- B. ネットワークトポロジーを検出し、ルーティングパスを計算し、スイッチ内の転送テーブル (IDテーブル) をプログラムします。
- C. ネットワークの輻輳を監視し、QoS (Quality of Service) メカニズムを使用してパケットの優先順位を動的に調整します。
- D. ユーザーが各InfiniBandスイッチ上でルーティングテーブルを手動で設定するためのコマンドラインインターフェイスを提供する。

E. ファイアウォールとして機能し、事前に定義されたルールに基づいて不正なトラフィックをブロックします。

正解: ([正解を表示します](#))

サブネットマネージャ (SM) は、InfiniBand トポロジの検出、経路の計算、およびスイッチ内の転送テーブル (LID テーブル) のプログラミングを担当します。これは、ファブリック内での接続確立と効率的なデータ転送の確保に不可欠です。InfiniBand は、IP アドレスを直接使用するのではなく、LID (ローカル識別子) ベースのルーティングを使用します。

有効的な **NCP-AII** 問題集は JPNTTest.com 提供され、**NCP-AII** 試験に合格することに役に立ちます！ JPNTTest.com は今最新 **NCP-AII** 試験問題集を提供します。JPNTTest.com NCP-AII 試験問題集はもう更新されました。ここで **NCP-AII** 問題集のテストエンジンを手に入れます。最新版のアクセス、<https://www.jpntest.com/shiken/NCP-AII-mondaishu> **125**問、**30%**ディスカウント、特別な割引コード: **JPNshiken**」

質問: 32

InfiniBand サーバーが動作を停止し、システム管理者が `ibstat` コマンドを実行したところ、以下の出力が得られました。

```
CA 'mlx5_1'  
CAタイプ :MT4115  
ポート数 2  
ファームウェアバージョン :10.20.1010  
ハードウェアバージョン: 0  
ノードGUID: 0x0002c90300002f78  
システムイメージGUID: 0x0002c90300002f7b  
ポート1 :  
状態: 初期化中  
物理状態: リンクアップ  
評価 :100  
ベース蓋 :0  
LMC: 0  
SM蓋 :0  
機能マスク: 0x0251086a  
ポートGUID: 0x0002c90300002f79  
リンク層: InfiniBand
```

この問題の原因は何ですか？

- A. HCAポートに不具合があります。
- B. 生地にランニングSMはありません。
- C. 隣接するスイッチポートに不具合があります。

D. ケーブルが切断されています。

正解: ([正解を表示します](#))

ibstatコマンドは、NVIDIA InfiniBandスタックにおける基本的な診断ツールであり、ローカルホストチャンネルアダプタ (HCA) の状態を照会するために使用されます。出力結果の中で最も重要なデータポイントは、物理状態、状態、およびSMリッドです。

物理状態: リンクアップは、サーバーの HCA と隣接するスイッチ ポート間の電氣的または光学的接続が物理層で確立され、正常であることを確認します。これにより、ケーブルの切断 (オプション D) またはハードウェア ポートの完全な故障 (オプション A および C) はすぐに除外されます。ただし、状態:

初期化中とは、「ワイヤ」は接続されているものの、論理的なInfiniBandプロトコルのハンドシェイクが完了していない状態を指します。

InfiniBandファブリックでは、サブネットマネージャ (SM) が、ノードの検出、ローカル識別子 (ID) の割り当て、ルーティングテーブルの設定を担当する中央集権型の「頭脳」です。出力に「Base lid: 0」および「SM lid: 0」と表示される場合、ポートにLIDが割り当てられておらず、通信可能なアクティブなサブネットマネージャが見つからないことを意味します。

ポートを「初期化中」から「アクティブ」に移行させるためのSMが実行されていない

と、RDMAトラフィックはファブリックを通過できません。このシナリオは通常、管理ノードでSMサービスがクラッシュした場合、または管理対象スイッチでSMが無効になっている場合に発生します。したがって、根本原因は、論理リンクの初期化を完了するための動作可能なサブネットマネージャがファブリック内に存在しないことです。

質問: 33

あなたはコロケーション施設に新しいAI推論サーバーを設置しようとしています。サーバーは、施設が管理する100GbEスイッチに接続されます。接続方法としては、LR4トランシーバーを使用したシングルモード光ファイバー接続、またはSR4トランシーバーを使用したマルチモード光ファイバー接続のいずれかを選択できます。サーバーとスイッチ間の距離は約75メートルです。コスト、信号品質、および将来の拡張性を考慮すると、どちらのオプションが最も適していますか？

A. LR4トランシーバーはシングルモードファイバーを使用しているため、長距離でもより優れた信号品質を提供します。

B. 短距離の場合、一般的にコスト効率が良いため、マルチモードファイバーを使用したSR4トランシーバー。

C. LR4トランシーバーはシングルモードファイバーを採用しており、電力効率が優れています。

D. どちらの選択肢も同等に適切であり、選択は任意です。

E. AOCケーブルは、他のどのタイプの接続よりも将来的な拡張性に優れているため、推奨します。

正解: ([正解を表示します](#))

75メートルの距離であれば、マルチモードファイバーを使用したSR4トランシーバーが最適です。通常、LR4トランシーバーよりもコスト効率が高く、この距離で十分な信号品質を提供します。LR4トランシーバーはより長距離向けに設計されており、価格も高くなります。AOCは便利ですが、マルチモードファイバーを使用したSR4の方がコスト効率に優れています。コスト効率こそが、両者を区別する決定的な要素です。

質問: 34

You're deploying a multi-GPU training job on a cluster using Slurm. You need to ensure that the GPUs allocated to the job are healthy and functioning correctly before the training starts. What's the MOST effective approach to pre-validate the GPU hardware?

- A.** Run a simple CUDA vector addition program on each GPU and check for errors.
- B.** Check the output of 'nvidia-smi' to ensure all GPUs are listed and have the expected memory.
- C.** Execute the NVIDIA Data Center GPU Manager (DCGM) diagnostic suite on the allocated GPUs.
- D.** Monitor the GPU temperature using 'nvidia-smi' during the first few minutes of the training job.
- E.** Allocate all available GPUs to the job and assume they are healthy.

正解: **C** ([コメントを发表する](#))

Using the DCGM diagnostic suite is the most thorough way to pre-validate GPU hardware. DCGM provides comprehensive tests to check GPU health, including memory, compute, and interconnects. A simple CUDA program or checking nvidia-smi provides basic validation, but not as comprehensive as DCGM. Monitoring temperature is reactive, not proactive. Assuming GPUs are healthy without validation is risky.

質問: 35

ある企業がNVIDIA DGX BasePODアーキテクチャを使用してAIファクトリーを導入しようとしています。インフラストラクチャチームは、コンピューティングノード間の高可用性と効率的なデータ転送を確保する必要があります。InfiniBandファブリックには、どのようなネットワークトポロジーを実装すべきでしょうか？

- A.** すべてのノードをループ状に接続するシンプルなリングトポロジー。
- B.** レール最適化設計のファットツリートポロジー。
- C.** すべてのトラフィックに対応する単一のフラットイーサネットネットワーク。
- D.** すべてのノードが単一の中央スイッチに接続されたスター型トポロジー。

正解: **B** ([コメントを发表する](#))

正解は、レール最適化設計のファットツリートポロジーです。NVIDIA DGX BasePOD および同様の AI Factory アーキテクチャは、高性能 InfiniBand ファブリックを使用して、分散 GPU トレーニング、低遅延通信、および計算ノード間のスケーラブルな帯域幅をサポートします。ファットツリートポロジーは、サーバーとスイッチ間に複数の高帯域幅パスを提供し、オーバーサブスクリプションの削減とバイセクション帯域幅の向上に役立ちます。

ルール最適化設計は、DGX 環境では特に重要です。各 GPU またはネットワークアダプタパスを特定のネットワークルールにアラインできるため、all-reduce や reduce-scatter などの NCCL コレクティブの予測可能な通信動作が向上します。単純なリングトポロジは、不要なホップ遅延と限られた障害耐性を導入するため、適していません。すべてのトラフィックに単一のフラット Ethernet ネットワークを使用しても、専用の高性能 InfiniBand ファブリックの要件を満たしません。中央に 1 つのスイッチがあるスタートポロジは、ボトルネックおよび単一障害点になる可能性があります。BasePOD規模のAIワークロードの場合、ファブリックは高可用性、帯域幅の一貫性、および効率的な東西GPU通信をサポートする必要があり、ルール最適化されたFat-Treeが適切な設計となります。

質問: 36

RDMAベースの分散トレーニングに使用されているサーバーで、NVIDIA ConnectX-6 ネットワークアダプタに不具合がある疑いがあります。アダプタのハードウェアと接続に関する潜在的な問題を診断するために、どのようなコマンドまたはツールを使用できますか？

- A. `lspci -v` を使用して、アダプタが検出され、そのリソースが正しく割り当てられていることを確認します。
- B. `ibstat` を使用して、アダプタの状態、リンク速度、アクティブなポートを確認します。
- C. `ethtool` を使用して、アダプタのイーサネット設定と統計情報を調べます。
- D. `ping` を実行して基本的なネットワーク接続をテストします。
- E. `nvsmimonitor` は GPU メトリックを監視し、異常を検出します。

正解: [\(正解を表示します\)](#)

E以外のすべてのオプションは、ネットワークアダプタの問題の診断に関連しています。

「`lspci -v`」 A) はハードウェア検出を確認します。「`ibstat`」 B) は InfiniBand 固有の詳細を確認します。「`ethtool`」 C) はイーサネット設定を調べます。「`ping`」 D) は基本的な接続性をテストします。「`nvsmimonitor`」 E) はネットワークアダプタではなく、GPU の監視に焦点を当てています。

質問: 37

ある金融サービス会社は、複数のサイトにわたる迅速な推論とデータ取得を必要とする、不正検出のためのAIモデルを導入しようとしています。同社のストレージシステムは、どの機能を優先すべきでしょうか？

- A. 中速で大容量。
- B. テープバックアップシステム。
- C. 低コストのHDDソリューション。
- D. 低遅延でのマルチプロトコルデータアクセス。

正解: [\(正解を表示します\)](#)

質問: 38

You are configuring an InfiniBand subnet with multiple switches. You need to ensure that traffic between two specific nodes always takes the shortest path, bypassing a potentially congested link. Which of the following approaches is MOST effective for achieving this using InfiniBand's routing capabilities?

- A.** Rely solely on the Subnet Manager's (SM) default path computation algorithm (e.g., Min Hop) without any modifications.
- B.** Use static routing by manually configuring forwarding tables on each switch along the desired path. This involves specifying DLID-to-Port mappings.
- C.** Implement Quality of Service (QOS) to prioritize the traffic between the two nodes, hoping that this will influence the path selection.
- D.** Utilize the `ibroute` command or similar tool to inject a static route between the nodes, forcing traffic to follow a specific path identified by LID and port number.
- E.** Decrease the MTIJ size on the potential congested link.

正解: ([正解を表示します](#))

Static routing with 'ibroute' (or similar) provides the most direct and reliable way to ensure traffic follows a specific path. The SM's default algorithm might not always choose the optimal path, and QOS only prioritizes traffic, not forces a specific route. Configuring forwarding tables manually on each switch is error-prone and difficult to manage at scale.

質問: 39

You're deploying BlueField OS to multiple SmartNICs with varying hardware revisions. How can you ensure that the correct device tree is loaded for each specific SmartNIC during the boot process?

- A.** Create a single device tree that is compatible with all SmartNIC revisions. The kernel will automatically handle compatibility.
- B.** ハードウェアのリビジョンを検出し、定義済みのマッピングに基づいて適切なデバイスツリーをロードできるブートローダー（例U-Boot）を使用します。
- C.** SmartNICごとにブートローダー構成でデバイスツリーファイルを手動で指定します
- D.** デバイスツリーをカーネルイメージ内に埋め込みます。カーネルは起動時に適切なものを自動的に選択します。
- E.** OSが起動した後、サービスが開始される前に、カスタムスクリプトを使用してデバイスツリーを再配置します。

正解: ([正解を表示します](#))

U-Bootのようなブートローダーは、ハードウェアの検出と、デバイスツリーなどのリソースの条件付きロードを処理するように設計されています。SmartNICのリビジョンを識別し、対応するDTBファイルをロードできます。互換性のある単一のDTBを作成することは難しく、ハードウェアの機能を十分に活用できない可能性があります。NICごとにDTBを手動で指定することは拡張性に欠けます。カーネルにDTBを埋め込むことは一般的ではありません。実行時にデバイスツリーを変更しようとする、不安定性を招く可能性があります。

質問: 40

After a recent OS upgrade, you need to reinstall NVIDIA GPU and DOCA drivers to support both AI training and accelerated networking. What best practice ensures successful installation and full hardware capability?

- A.** Download and install only the specific versions of GPU and DOCA drivers listed as compatible with the current OS and hardware.
- B.** Apply legacy drivers for hardware released within the last two years to maintain maximum compatibility across versions.
- C.** Install the latest available drivers directly from the NVIDIA website.
- D.** Use the default drivers provided by the Linux distribution, unless an installation fails during system boot.

正解: ([正解を表示します](#))

The correct best practice is to install only the GPU and DOCA driver versions that are validated for the current operating system, kernel, hardware platform, and NVIDIA software release. In NVIDIA AI infrastructure, driver compatibility is critical because GPU drivers, CUDA libraries, DCGM, Fabric Manager, DOCA-OFED, network adapters, DPUs, and kernel modules must align. Installing the newest driver without checking compatibility can cause driver/library mismatches, failed nvidia-smi, broken RDMA, or unsupported DOCA behavior. NVIDIA DGX documentation notes that DGX OS updates include coordinated OS, kernel, GPU driver, CUDA toolkit, and DCGM updates, and NVIDIA AI Enterprise provides support matrices to confirm supported platform combinations. DOCA-OFED is also tied to the operating system and kernel environment, so version selection must be deliberate. Legacy drivers may lack required fixes or hardware support, and default distribution drivers are usually not sufficient for validated DGX or AI Enterprise deployments. After installation, administrators should reboot if required, validate GPU visibility, confirm networking driver status, and run health checks before returning the node to production.

質問: 41

After ClusterKit reports " GPU-Host latency exceeds threshold, " which NVIDIA diagnostic tool should be used to isolate hardware faults?

- A.** Re-run ClusterKit with --stress=gpu -Y 60 to extend test duration
- B.** nvidia-smi topo -m to inspect GPU topology connections
- C.** DCGM Diags dcgmi diag -r 2
- D.** ib_write_bw to measure InfiniBand bandwidth between nodes

正解: **B** ([コメントを発表する](#))

" GPU-Host latency " issues in NVIDIA DGX or HGX systems are frequently caused by incorrect PCIe affinity or sub-optimal NUMA (Non-Uniform Memory Access) mapping. If a GPU is forced to communicate with a CPU core or an HCA that is not on its local PCIe switch/root complex, latency increases significantly as data must cross the QPI/UPI inter-

processor links. The command `nvidia-smi topo -m` provides a detailed matrix of the system's internal topology, showing how GPUs, CPUs, and NICs are connected. It identifies whether the connection is via a single PCIe switch (PIX), multiple switches (PXB), or across the CPU (SYS).

By inspecting this map, an administrator can identify if a software process is pinned to the wrong NUMA node or if a hardware path is unexpectedly degraded. While DCGM (Option C) is good for checking component health, it doesn't map the logical-to-physical affinity paths that cause specific latency " threshold " warnings.

質問: 42

Kubernetes クラスタ内に、異なるバージョンの NVIDIA ドライバを実行しているノードが複数存在します。コンテナ化された AI アプリケーションが、実行されるノードで実行されている特定のドライババージョンと常に互換性があることを保証する必要があります。クラウドネイティブな方法で、このドライババージョンの互換性をどのように実現できますか？

- A. ドライバのバージョンごとに異なるコンテナイメージを手動で作成し、ノードセレクタを使用して適切なイメージを適切なノードにスケジュールします。
- B. NVIDIA ドライバの機能を使用して、実行時にドライバーのバージョンを検出し、適切なライブラリを動的にロードします。
- C. NVIDIA Operator を使用して、ノード上のドライバのインストールと更新を自動的に管理し、クラスタ全体で一貫したドライババージョンを確保します。
- D. ノードラベルを検査し、実行時に適切な NVIDIA ライブラリをポッドに注入する Webhook を実装します。
- E. 共有ボリュームを使用してドライバをコンテナにマウントします。

正解: ([正解を表示します](#))

NVIDIA Operator は、Kubernetes クラスタ内で NVIDIA ドライバおよび関連コンポーネントのライフサイクルを管理するように設計されており、自動インストール、アップデート、バージョン管理などが含まれます。これにより、すべてのノードで一貫性のある互換性のあるドライババージョンが確保され、アプリケーションのデプロイと管理が簡素化されます。Operator は、NVIDIA スタックを管理するためのクラウドネイティブなアプローチです。オプション A は手動であり、柔軟性に欠けます。オプション B は、ホスト上での NVIDIA ドライバの適切なセットアップに依存します。オプション D は複雑な作業です。オプション E は安全ではありません。

質問: 43

システム管理者は、NVIDIA DGX A100 に潜在的なハードウェア障害が発生したというアラートを受け取りました。GPU のパフォーマンスが低下し、システムファンが大きな音で動作しています。ハードウェア障害を特定してトラブルシューティングするために、どのような手順を推奨すべきでしょうか？

- A. ディープラーニングのワークロードを実行してGPUに負荷をかけ、問題が解消されるかどうかを確認してください。
- B. NVIDIAシステム管理インターフェイス (nvidia-smi) でGPUの状態と温度を確認してください。
- C. 電源を切ってからDGXを再起動し、パフォーマンスの低下が解消されるかどうかを確認します。
- D. ファン速度を最大まで上げて、パフォーマンスが向上するかどうかを確認してください。

正解: **B** ([コメントを发表する](#))

DGX システムでファン速度が速くなり、パフォーマンスが低下する場合、通常はサーマルスロットリングが発生しています。A100 や H100 などの高性能 GPU は、安全な温度しきい値を超えると、クロック速度 (したがってパフォーマンス) を自動的に低下させます。最初に行うべき最も重要な診断手順は、nvidia-smi を実行することです。このユーティリティは、GPU の温度、消費電力、および「クロック スロットリングの理由」に関するリアルタイムのテレメトリを即座に提供します。出力を確認することで、管理者はクロック低下の理由として「熱」がリストされているかどうかを確認できます。これにより、問題が環境 (エアフローの遮断/ホットアイルの温度) によるものか、ハードウェア固有のもの (GPU サーマルインターフェイスの故障または内部ファンの故障) によるものかを特定できます。ワークロードを増やす (オプション A) と熱が悪化しますが、電力消費 (オプション C) は診断データを提供しない「最終手段」です。nvidia-smi は、GPU トレイの RMA (返品承認) が必要かどうかを判断するために必要な証拠データを提供します。

質問: **44**

Redfish経由でBlueField-3 DPU BMCファームウェアをアップデートした後、エンジニアは15分経過しても「TaskState: Running」と表示されるものの、進捗がないことを確認しました。アップデートの完了状況をどのように追跡すればよいのでしょうか？

- A. DPUオペレーティングシステムの/var/log/messagesで更新ログを確認してください。
- B. インストールプロセスのタスクIDを使用してDPU BMCにクエリを実行します。
- C. DPUを直ちに再起動して、強制的にロールバックしてください。
- D. DPU上でbfrec --statusを実行してフラッシュの進行状況を確認します。

正解: ([正解を表示します](#))

正しい対処方法は、Redfishファームウェア更新プロセスによって作成されたタスクIDを使用してDPU BMCに問い合わせることです。Redfishを介して実行されるBlueField BMCファームウェア更新は、BMC側のタスクとして追跡されます。

NVIDIA BlueField BMC firmware documentation describes Redfish-based firmware upgrade operations and indicates that the secure update starts tracking progress. The Redfish task object exposes fields such as TaskState, TaskStatus, messages, and progress indicators, which are the correct source of truth for update completion or failure. Checking /var/log/messages inside the DPU operating system may not show BMC

firmware update progress because the operation is managed by the BMC. Power cycling immediately is unsafe because interrupting firmware update or activation can cause recovery problems. `bfrec --status` is not the standard command for tracking Redfish BMC firmware update task progress. In NVIDIA AI infrastructure bring-up and maintenance, firmware operations should be monitored through the management controller workflow that initiated them, preserving the Task ID and polling the BMC task endpoint until completion, failure, or a documented timeout condition is reached.

質問: 45

出力結果を参照してください。

```
~ $ sudo nvsm show healthinfo
```

```
タイムスタンプ: 2017年12月16日(土) 16:26:32 -0800
```

```
バージョン: 17.12-5
```

```
BIOSリビジョンチェック[5.11].....
```

```
DGXシリアル番号[YSY72800016].....
```

```
インストールされているDIMMメモリスティックを確認します.....正常です
```

```
...[出力は省略されています]
```

```
イーサネットコントローラの検証.....正常
```

```
インストールされているGPUを確認してください。.....正常ではありません
```

```
想定されるGPUについて「lspci」の出力を確認しています
```

```
PCIアドレス「07:00.0」にGPUが見つかりません
```

```
インストールされているInfiniBandコントローラを確認します.....正常
```

```
PCIeスイッチの検証.....正常
```

```
...[出力は省略されています]
```

システム管理者は、DGXシステムの健全性に関してどのような知見を得ることができるでしょうか？

- A. GPUトレイのアップグレードに失敗しました。
- B. DGXシステムにGPUがありません。
- C. GPUドライバのアップグレードに失敗しました。
- D. システムはハードウェアの健全性チェックに正常に合格しました。

正解: [\(正解を表示します\)](#)

出力結果は、NVIDIAシステム管理 (NVSM) ツール、具体的には `nvsm show healthinfo` コマンドの実行結果です。NVSMは、NVIDIA DGXシステムにとって不可欠な診断フレームワークであり、ハードウェアの状態を監視し、障害を特定し、システムが検証済みの動作状態を維持するのに役立ちます。

この特定の診断トレースでは、システムは「インストールされているGPUの検証」チェックが「異常」ステータスを返したと報告しています。根本原因を特定するために、NVSMは `lspci` コマンドからのライブハードウェア列挙を、システムの既知の「ゴールド構成」ウェアームウェアで定義されているハードウェアマニフェスト)と照合します。明示的なエラーメッ

「セージ PCIアドレス'07:00.0'にGPUがありません」は、システムがその特定のPCIeバスアドレスにGPUモジュールが存在することを期待しているにもかかわらず、ハードウェアが応答していないか、バスから認識されていないことを示しています。

この情報により、システム管理者は、システムの論理的な観点から、GPUが欠落していると結論付けることができます。これは、ソフトウェアやドライバの問題ではなく、重大なハードウェア障害です。DGX H100またはA100システムでは、物理モジュールの故障、GPUベースボードの特定のセグメントへの電力供給の問題、またはPCIeスイッチファブリックの障害が原因である可能性があります。DGXは高速集合通信 (NCCL)に8つのGPUすべてを必要とするため、GPUが1つでも欠落すると、ノードは大規模なトレーニングジョブに参加できなくなり、物理的な検査またはGPUトレイの交換 (RMA)が必要になります。

質問: 46

クラスター拡張後、リーフスイッチでトランシーバーの「ファームウェアバージョン不一致」アラートが表示されます。トランシーバーのファームウェアを想定されるバージョンと比較して検証するツールはどれですか？

- A. フリント
- B. iblinkinfo
- C. mlxconfig
- D. ethtool

正解: ([正解を表示します](#))

ファームウェアの一貫性は、InfiniBandファブリックの安定したパフォーマンスの基盤となります。クラスターを拡張すると、新しいトランシーバーやケーブルに既存のベースよりも新しいファームウェアや古いファームウェアが搭載されている場合があります、UFM (Unified Fabric Manager)などの管理コンソールで「FWバージョン不一致」アラートが発生する可能性があります。flintツール (またはmstflint)は、トランシーバーに組み込まれている特定のファームウェアレベルを照会するための適切なユーティリティです。iblinkinfoはリンク速度やポートの状態に関するデータを提供しますが、バージョン検証に必要なハードウェアレベルの詳細なファームウェアテレメトリは提供しません。flintを使用すると、管理者はデバイスを照会し、現在の書き込みバージョンをターゲットイメージと比較し、クラスターを統一された状態にするために必要な更新を実行できます。NVIDIA AIインフラストラクチャでは、ファブリック全体でファームウェアを統一することで、アダプティブルーティングや輻輳制御などの機能が予測どおりに動作することが保証されます。バージョンが一致していない場合、前方誤り訂正 (FEC)やリンクアップネゴシエーションにおける動作の不整合により、アプリケーション (NCCL)レベルでの診断が困難な断続的なパフォーマンス低下が発生する可能性があります。

有効的な**NCP-AII**問題集はJPNTTest.com提供され、**NCP-AII**試験に合格することに役に立ちます！JPNTTest.comは今最新**NCP-AII**試験問題集を提供します。JPNTTest.com NCP-AII試験問題集はもう更新されました。ここで**NCP-AII**問題集のテストエンジンを手に入れます。最新版のアクセス、<https://www.jpntest.com/shiken/NCP-AII-mondaishu>
「125問、30%ディスカウント、特別な割引コード: **JPNshiken**」

質問: 47

AIトレーニング用に設計されたサーバーに、NVIDIA AIOO GPUを4枚インストールしようとしています。サーバーのマザーボードには複数のPCIe Gen4 x16スロットがありますが、サーバーの電源ユニット (PSU)には8ピンPCIe電源コネクタが3つしかありません。すべてのGPUに十分な電力を供給するための最善策は何でしょうか？

- A. 8ピンコネクタの1つにPCIe電源スプリッターケーブルを使用して、2つのGPUに電源を供給します。
- B. GPUを3枚だけインストールし、4枚目は電源を入れずにします。
- C. GPUのクロック周波数を大幅に下げて、利用可能な電源容量よりも消費電力を削減します。
- D. 既存の電源ユニットを、少なくとも4つの8ピンPCIe電源コネクタを備えた、より高出力の電源ユニットに交換してください。
- E. マザーボードの内部SATA電源コネクタを使用してGPUを接続します。

正解: ([正解を表示します](#))

スプリッターを使用すると、元のコネクタに過負荷がかかり、不安定になったり損傷したりする可能性があります。GPUに電源を供給しないのは本末転倒です。SATAコネクタはGPUの高い電力要件に対応するように設計されていません。クロック周波数を下げることでも解決できる場合もありますが、長期的な解決策としては信頼できません。電源ユニットを交換するのが最も安全で確実な解決策です。

質問: 48

NVIDIA認定プロフェッショナルA1インフラストラクチャ (NCP-AII) 環境で、ホスト用のネットワークファブリックポートを設定しています。ネットワークスループットを向上させるために、ジャンボフレームを設定する必要があります。ネットワークインターフェイスとスイッチに設定する一般的なMTU (最大伝送単位サイズはどれくらいですか？また、その理由は？

- A. 1500バイト。これはデフォルトであり、ほとんどのネットワークと互換性があります。
- B. 9000バイト、別名ジャンボフレームは、オーバーヘッドを削減し、大容量データ転送のスループットを向上させます。
- C. 65535バイト。理論上の最大MTUサイズで、最高のパフォーマンスを実現します。
- D. 576バイト、IPv4で必要とされる最小MTUサイズ。
- E. MTUサイズは1500~9000バイトの間で指定できます。具体的な値は問いません。

正解: ([正解を表示します](#))

A1訓練環境で最適なパフォーマンスを得るには、ジャンボフレーム (MTU 9000バイト)の使用を推奨します。MTUを大きくすることで、一定量のデータを送信するために必要なパケット数を減らし、処理オーバーヘッドを削減して全体的なスループットを向上させます。技術的にはより大きなMTUも可能ですが、ジャンボフレームの標準は9000バイトです。

質問: 49

DockerコンテナでNVIDIA GPUを使用することに関して、以下の記述のうち正しいものはどれですか？

- A.** NVIDIA Container Toolkitを使用すると、コンテナイメージを変更することなく、GPUアクセラレーション対応アプリケーションをDockerコンテナ内で実行できます。
- B.** GPUサポートを有効にするには、Dockerコンテナ内にNVIDIAドライバーをインストールする必要があります。
- C.** 'nvidia-smr' コマンドはホストマシンでのみ実行でき、Docker コンテナ内では実行できません。
- D.** アプリケーションが CODA を使用する場合、コンテナ内に CUDA ライブラリが必要です。
- E.** コンテナ内で 'CUDA_VISIBLE_DEVICES'などの環境変数を使用すると、アプリケーションがアクセスできるGPUに影響を与えることができます。

正解: **A,D,E** ([コメントを发表する](#))

NVIDIA Container Toolkit を使用すると、イメージを変更することなく、GPU アクセラレーション対応アプリを Docker で実行できます。ホストのドライバが活用されます。アプリが CUDA を使用する場合は、コンテナ内に CUDA ライブラリが必要です。は、コンテナ内の GPU の可視性を制御するために使用されます。ドライバはホストによって管理されるため、コンテナ内には不要です (B は誤りです)。また、NVIDIA Container Toolkit が正しくセットアップされていれば、コンテナ内で 'nvidia-smi' を実行できます (C は誤りです)。

質問: 50

An infrastructure engineer is preparing a new AI cluster for production use, relying on NVIDIA switches and high-speed optical transceivers for node connectivity. The team is finalizing network validation before launching large-scale training jobs. Why is it critical to confirm and align the firmware version on all switch transceivers prior to production?

- A.** To guarantee that hardware inventory tools can report serial numbers and manufacturer codes for asset management, which is critical for future support and troubleshooting.
- B.** To ensure stability, bandwidth, and compatibility across the cluster, avoiding link issues and performance loss.
- C.** To allow the network operating system to automatically discover all connected transceivers with heterogeneous firmware.
- D.** To reduce GPU memory consumption during distributed training jobs.

正解: ([正解を表示します](#))

正しい答えは、クラスタ全体で安定性、帯域幅、および互換性を確保することです。高速な NVIDIA AI ファブリックでは、光トランシーバーとアクティブケーブルは単なる受動的な要素ではなく、リンクトレーニング、信号品質動作、ファームウェア制御動作、診断、およびスイッチやアダプタとの互換性に関与します。ファブリック全体でトランシーバーのファームウェアバージョンが一致しない場合、リンクのネゴシエーションが異なったり、テレメトリが不整合になったり、断続的なエラーが発生したり、持続的なワークロードでパフォーマンスが低下したりする可能性があります。これは、NCCL コレクティブがレイテンシ、再送信、および帯域幅の変動に敏感であるため、AI トレーニングにとって非常に重要です。インベントリレポートは有用ですが、ファームウェアを整合させる主な理由ではありません。

異種混在のトランシーバーファームウェアは、検出のためだけに望ましいものではなく、トラブルシューティングを困難にする可能性があります。GPU メモリ消費量は、スイッチトランシーバーファームウェアとは無関係です。物理層検証の際には、エンジニアはファブリックを本番環境向けに承認する前に、サポートされているトランシーバーモデル、ファームウェアバージョン、リンク速度、エラーカウンタ、BER 健全性、ポート安定性を確認する必要があります。ファームウェアの整合性により、AI クラスタ全体で予測可能でサポート可能なベースラインを作成できます。

質問: 51

You have installed NVIDIA drivers using the '.run' installer on a system running Ubuntu. However, after each kernel update, the NVIDIA drivers stop working. What is the most effective way to address this issue permanently?

- A. Re-run the '.run' installer after each kernel update.
- B. Use DKMS (Dynamic Kernel Module Support) to automatically rebuild the NVIDIA kernel modules after each kernel update.
- C. Disable automatic kernel updates.
- D. Manually compile the NVIDIA kernel modules after each kernel update.
- E. Create a script that symlinks the old kernel modules to the new kernel directory.

正解: **B** ([コメントを发表する](#))

DKMS (Dynamic Kernel Module Support) が最も効果的な解決策です。新しいカーネルがインストールされるたびに、NVIDIA などのカーネルモジュールを自動的に再構築します。Irun インストーラーを再実行する作業は手動で、エラーが発生しやすいです。セキュリティ上の理由から、カーネルのアップデートを無効にすることは推奨されません。手動でコンパイルするのは複雑で時間がかかります。カーネル API の変更により、古いモジュールをシンボリックリンクしても機能しない可能性が高いです。

質問: 52

あなたは、複数の NVIDIA A100 GPU を搭載した AI インフラストラクチャを管理しており、各 MIG インスタンスの使用状況を個別に監視したいと考えています。MIG インスタンスレ

ベルの詳細な監視データを提供するツールはどれですか？（該当するものをすべて選択してください）

- A. nvidia-smi
- B. DCGM (データセンターGPUマネージャー)
- C. gpustat
- D. htop
- E. nvidia-docker

正解: **A,B** ([コメントを发表する](#))

'nvidia-smi と DCGM (Data Center GPU Manager) は、MIG インスタンスを含む NVIDIA GPU を詳細レベルで監視するための主要なツールです。'nvidia-smi は、個々の MIG インスタンスの使用率、メモリ使用量、その他のメトリックを表示できます。DCGM は、より広範なメトリックとアラート機能を備えた、より包括的な監視ソリューションを提供します。gpustat は GPU 統計情報を提供しますが、'nvidia-smi や DCGM ほど詳細な MIG レベルの情報は提供しない場合があります。'htop はシステムレベルのプロセス モニターであり、GPU 固有のメトリックは提供しません。'nvidia-docker は、監視ではなく、NVIDIA GPU を使用したコンテナ化に関連しています。

質問: **53**

You encounter a situation where a container running with GPU support is experiencing significant performance degradation compared to running the same application directly on the host. You have already verified that the NVIDIA drivers are correctly installed and the NVIDIA Container Toolkit is properly configured. Which of the following could be contributing factors to this performance difference?

(Select all that apply)

- A. The container is using a significantly older version of the CUDA runtime compared to the host.
- B. CPU pinning or NUMA affinity is not properly configured for the container, leading to inefficient memory access.
- C. The '-ipc=host' flag is not used when running the container, causing inter-process communication overhead.
- D. The kernel version within the container is significantly different from the host kernel, leading to driver compatibility issues.
- E. Insufficient bandwidth between CPU and GPU

正解: ([正解を表示します](#))

コンテナ内で古い CUDA ランタイム (A) を使用すると、最適化が不足していたり、アプリケーションとの互換性の問題により、パフォーマンスが低下する可能性があります。不適切な CPU ピンニングと NUMA アフィニティ (B) は、特にマルチソケット システムにおいて、コンテナがメモリに非効率的にアクセスする原因となる可能性があります。'-ipc=host' (C) は、ホストの IPC 名前空間を共有することで、場合によってはパフォーマンスを向上させることができますが、常に必要というわけではなく、セキュリティ上の問題

が発生する可能性があります。カーネルバージョンの違い (D) は通常、互換性を保証する NVIDIA Container Toolkit によって処理されます。CPU と GPU 間の帯域幅不足 (E) は、ハードウェアの問題が原因である可能性があります。

質問: 54

導入後のチェックで、NVIDIA GPUアクセラレーション対応データセンターの環境を検証しています。NVIDIA DGX H100システムの安定動作に必要な電力と冷却が確保されていることを確認するために、どの操作が不可欠ですか？

- A. 過熱を防ぐため、すべての負荷条件下でシステムファンが100%で動作していることを確認してください。
- B. システムBIOSを確認し、GPUオーバークロックが有効になっていて最高のパフォーマンスが得られることを確認してください。
- C. NVSMを使用して未使用のPCIeデバイスを無効にし、システム全体の発熱量を削減します。
- D. 各DGXシステムが冗長構成で適切な定格のPDUに接続されており、すべての電源装置が公称入力を報告していることを確認します。

正解: D ([コメントを发表する](#))

DGX H100のような高密度AIインフラストラクチャの安定動作には、電力および熱仕様への厳格な準拠が必要です。単一のDGX H100システムは、ピーク負荷時に最大10.2kWの電力を消費する可能性があります。そのため、最も重要な検証手順は、電気的なインフラストラクチャからサーバーへの「引き渡し」が正常であることを確認することです。これには、システムが、ブレーカーをトリップさせることなく電流要件を処理できる冗長PDU（電源分配ユニット）に接続されていることを確認することが含まれます。NVSM (NVIDIAシステム管理) を使用して、管理者は6つの電源ユニット (PSU) すべてが機能しており、公称入力電圧（通常00V~240V）を受信していることを確認する必要があります。PSUが最適でない入力または「冗長性の喪失」を報告すると、負荷の高いトレーニング実行中にシステムがパフォーマンスを低下させたり、予期せずシャットダウンしたりする可能性があります。常に100% (オプションA) で動作させるのは、冷却ポリシーが非効率的または失敗していることを示しています。ファンは温度に応じて動的に調整されるべきだからです。オーバークロック (オプションB) は、エンタープライズ向けDGXシステムではサポートされておらず、推奨もされていません。これらのシステムは、すでに工場出荷時に最高の安定性能を発揮するように調整されているためです。

質問: 55

DGXノードで24時間のストレステストを実行した後、システム管理者はシステムの安定性を確保するために、どの2つの重要な指標を確認すべきでしょうか？

- A. 平均CPU使用率が80%以上で、Dockerコンテナの稼働時間も良好であること。
- B. テスト全体を通して、サーマルスロットリングは発生せず、GPU使用率は95%以上を維持しました。

- C. SSDの書き込み耐久性とRAM容量。
- D. 総エネルギー消費量とNVLink帯域幅。

正解: ([正解を表示します](#))

HPLやNCCLなどのツールを使用した24時間のストレステストは、DGXシステムの熱的および電氣的な限界を押し上げるように設計されています。合格を確認するには、管理者はハードウェアが性能目標を劣化なく維持したことを確認する必要があります。GPU使用率が95%以上で安定していることは、ワークロードがテスト期間全体にわたって計算コアを正常に飽和させたことを示しています。重要なのは、サーマルスロットリングイベントがない (nvidia-smi -q -d PERFORMANCEで確認) ことで、システムの冷却ソリューション (ファンとヒートシンク) が環境に適切であることを確認することです。スロットリングが発生した場合、GPUは自身を保護するために速度を落とし、冷却の故障または環境の熱問題の可能性を示します。消費電力 (オプションD) とCPU使用率 (オプションA) は興味深いものですが、極端なAIトレーニング負荷下での「安定性」の主要な指標ではありません。システムの安定性は、ハードウェアレベルの介入や速度低下なしに、ピーク速度で無期限に実行できる能力によって定義されます。

質問: 56

あなたはNVIDIA NVLinkスイッチシステムの構成を担当しています。GPUとスイッチを物理的に接続した後、ソフトウェア構成プロセスにおける一般的な最初のステップは何ですか？

- A. 接続されているすべてのGPUに最新のNVIDIAドライバーをインストールします。
- B. システムBIOSを設定してNVLinkサポートを有効にする。
- C. NVLinkスイッチのファームウェアを更新しています。
- D. NVLinkスイッチ管理ソフトウェアをインストールします。
- E. 接続されているすべてのGPU間でメモリ帯域幅テストを実行しています。

正解: ([正解を表示します](#))

Updating the NVLink Switch firmware is crucial for ensuring compatibility and stability with the connected GPUs and the overall system. It addresses potential bugs, security vulnerabilities, and performance issues. It should always be done first before any other software configuration. BIOS settings should be checked beforehand, and the NVLink management software comes after the firmware update.

質問: 57

A security policy requires you to log all NGC CLI commands executed on a specific host. How can you achieve this without modifying the NGC CLI source code?

- A. Use the 'history' command in Linux to view the command history for the user account used to execute the NGC CLI commands.
- B. Implement system-level auditing using tools like 'auditd' (Linux) or Windows Event Logging to capture all executed commands, including NGC CLI commands.

C. Create a wrapper script around the 'ngc' command that logs the command and its arguments before executing the actual NGC CLI command.

D. NGC CLI automatically logs all commands to a file in the '/var/log/ngc' directory.

E. Configure the NGC CLI to send all commands to a remote syslog server.

正解: [\(正解を表示します\)](#)

System-level auditing (B) provides a comprehensive way to log all executed commands.

Creating a wrapper script (C) allows you to log the command before it's executed. The 'history' command (A) only captures commands executed in the current shell session.

Option D is incorrect as NGC CLI does not log automatically to this directory, and option E is not a supported feature.

質問: 58

クラスタ展開中に、UFMケーブル検証ツールが複数のInfiniBandリンクで「誤ったネイバー」エラーを報告します。この問題を解決する最も効率的な方法は何ですか？

A. すべてのリーフスイッチを再起動して、LLDPの再検出を強制します。

B. 影響を受けたケーブルをすべて、より高グレードのOM5光ファイバーに交換します。

C. LLDP データをトポロジ ファイルと照合して検証し、修正します。

D. 近隣検証をバイパスするために、すべてのスイッチでFECを無効にします。

正解: [C \(コメントを发表する\)](#)

NVIDIA DGX SuperPOD などの大規模な InfiniBand ファブリックでは、アダプティブルーティングおよびファット ツリー アルゴリズムが正しく機能するために、ケーブル トポロジを正確に維持することが必須です。Unified Fabric Manager (UFM) が、ケーブルがマスター トポロジ マップ (多くの場合 .csv または .topology ファイル) で指定されたポート以外のポートに接続されていることを検出した場合、「間違ったネイバー」エラーが発生します。UFM は、LLDP (リンク レイヤ ディスカバリ プロトコル) またはサブネット管理パケットを使用して、リンクの両端の GUID を識別します。

最も効率的な修復方法は、UFMから提供されるライブLLDPデータと設計図を照合することです。これにより、エンジニアはエラーが物理的なケーブル接続ミス (ポートの入れ替え) によるものか、トポロジーファイルの論理エラーによるものかを特定できます。スイッチの再起動 (オプションA) では物理的なパッチエラーは修正されず、FECの無効化 (オプションD) は根本的なルーティングロジックの問題に対処しないまま、400G (NDR) リンクで壊滅的な信号損失を引き起こします。物理的なパッチを修正するか、トポロジーファイルを更新することで、ファブリックの「グラウンドトゥールズ」が確実に復元されます。

質問: 59

高性能ストレージ向けにBlueField-3 DPUベースのサーバーを構成しています。リモートのNVMe SSDにアクセスするためにNVMe-oF (NVMe over Fabrics) を使用したいと考えています。このNVMe-oF構成において、BlueField DPUを使用する主な利点は、標準NICを備えた従来のサーバーと比較して何でしょうか？

- A. BlueField DPUは、手動操作なしでNVMe-oFターゲットを自動的に構成します。
- B. BlueField DPUはNVMe-oFプロトコル処理をオフロードし、ホストサーバーのCPUオーバーヘッドを削減します。
- C. BlueField DPUは、別途NVMe-oFターゲットサーバーを必要としません。
- D. BlueField DPUは、すべてのNVMe-oFトラフィックに対してハードウェア暗号化を内蔵しています。
- E. BlueField DPUは、NVMe-oF接続を中断することなくNVMe SSDのホットスワップを可能にします。

正解: ([正解を表示します](#))

BlueField DPUをNVMe-oF構成で使用する最大の利点は、NVMe-oFプロトコル処理をオフロードできることです。これにより、ホストサーバーのCPU負荷が大幅に軽減され、他のタスクにより多くのリソースを割り当てることが可能になります。一部のDPUはハードウェア暗号化などの機能を提供する場合がありますが、主なメリットはプロトコル処理のオフロードです。

質問: 60

NVIDIA A100 GPU 上に 2 つの lg.10gb MIG インスタンスを構成しました。一方のインスタンスでディープラーニングのトレーニングジョブを実行しており、もう一方の MIG インスタンスのリソースを消費しないようにしたいと考えています。ハードウェア レベルで 2 つの MIG インスタンス間の分離を保証するメカニズムはどれですか？

- A. CUDAマルチプロセスサービス (MPS)
- B. vGPUスケジューリング
- C. MIG ハードウェア パーティショニング
- D. Kubernetesのリソース割り当て
- E. cgroups

正解: C ([コメントを发表する](#))

MIG (Multi-Instance GPU) provides hardware-level partitioning of the GPU. This means that each MIG instance has dedicated compute, memory, and memory bandwidth resources, ensuring strong isolation between the instances. CUDA MPS allows multiple processes to share a single GPU, but does not provide isolation. vGPU scheduling is for virtualized environments. Kubernetes resource quotas provide resource limits at the container orchestration level but do not provide hardware-level isolation. Cgroups is a Linux kernel feature to limit, account, and isolate resource usage (CPU, memory, disk I/O, network, etc.) of process groups. It is often used in conjunction with containerization technologies such as Docker, but does not directly influence the MIG hardware partitioning.

質問: 61

A system administrator needs to configure a BlueField DPU and enable RShim on the baseboard management controller (BMC). Which command should be executed?

- A. `ipmitool raw 0x32 0x6a 1`
- B. `systemctl restart rshim`
- C. `systemctl enable bmc-rshim.service`
- D. `scp < path_to_bfb > root@ < bmc_ip > :/dev/rshim0/boot`

正解: ([正解を表示します](#))

In NVIDIA BlueField DPU architectures, the RShim (Remote-Shim) interface provides a vital communication channel between the DPU and the host or BMC, typically used for early-stage provisioning, console access, and firmware loading. While the DPU is usually managed via the host 's PCIe bus, certain data center configurations require the DPU to be managed out-of-band via the server 's Baseboard Management Controller (BMC). To enable this capability, a specific low-level command must be sent to the BMC to toggle the RShim functionality over the internal USB-to-BMC bridge. The command `ipmitool raw 0x32 0x6a 1` is the verified raw IPMI hex code used in NVIDIA DGX and certified systems to enable the BMC-to-DPU RShim path. Once enabled, the BMC can " see " the DPU as a USB device, allowing the administrator to push a BlueField Boot (BFB) image to `/dev/rshim0/boot` for OS installation even if the host CPU is powered off or unresponsive. Option B and C are host-side service commands that assume the driver is already loaded and the hardware path is active, whereas the raw IPMI command is required to enable the hardware path itself.

有効的な**NCP-AII**問題集はJPNTTest.com提供され、**NCP-AII**試験に合格することに役に立ちます！JPNTTest.comは今最新**NCP-AII**試験問題集を提供します。JPNTTest.com NCP-AII試験問題集はもう更新されました。ここで**NCP-AII**問題集のテストエンジンを手に入れます。最新版のアクセス、<https://www.jpntest.com/shiken/NCP-AII-mondaishu>
125問、30%ディスカウント、特別な割引コード: **JPNshiken**」

質問: 62

An engineer must ensure that a BlueField-3 NIC firmware download matches the cluster's PSID. Which step is critical before installation?

- A. Check that the DPU's BMC IP is reachable by ping.
- B. Confirm that the firmware file size matches the DPU's flash capacity.
- C. Use `mstflint -d < PCI_ID > query` to validate the device PSID before selecting the firmware image.
- D. Verify that the SHA256 hash of the firmware matches NVIDIA's public ledger.

正解: ([正解を表示します](#))

The critical step is to query the device PSID using `mstflint -d < PCI_ID > query` before selecting or installing the firmware image. The PSID identifies the exact NVIDIA networking device variant, board configuration, and supported firmware branch. In a BlueField-3 or

ConnectX-based AI infrastructure environment, using firmware built for the wrong PSID can cause installation failure, unsupported behavior, or device instability.

This is especially important in cluster operations because firmware consistency must be maintained across nodes, DPUs, SuperNICs, and adapters. Checking BMC reachability may help with platform management but does not confirm firmware-device compatibility. Firmware file size is not a reliable compatibility indicator.

Hash verification can confirm file integrity, but it does not prove the image matches the device PSID.

NVIDIA MFT and mstflint workflows are designed for querying device identity and firmware details before burn or update operations, and NVIDIA documentation references PSID-based firmware querying in MFT workflows.

質問: 63

Dockerコンテナ内で実行されているGPUアクセラレーション対応アプリケーションのパフォーマンス問題のトラブルシューティングを行っています。コンテナ内の `nvidia-smi` の出力ではGPUが使用されていることが示されていますが、パフォーマンスは予想よりも大幅に低くなっています。このパフォーマンスボトルネックの原因として考えられるのは次のうちどれですか？

- A. ホストマシンのCPUが過剰に使用されているため、GPUへのデータ転送にボトルネックが発生しています。
- B. Dockerコンテナは、GPUとのデータ転送に共有メモリを使用するように構成されていません。
- C. ホスト上の CUDA ドライバのバージョンが、コンテナで使用されている CUDA ツールキットのバージョンと互換性がありません。
- D. アプリケーションは、CPU と GPU 間で頻繁に小さなメモリ転送を実行します。
- E. GPUが過熱し、サーマルスロットリングが発生しています。

正解: ([正解を表示します](#))

GPUが使用されている場合でも、Dockerコンテナ内でGPUのパフォーマンスが低下する原因はいくつか考えられます。CPUの負荷が高い場合 A)、GPUへのデータ転送がボトルネックになる可能性があります。ホストとコンテナ間でCUDAドライバのバージョンが互換性がない場合 C)、予期しないエラーが発生する可能性があります。CUDAドライバはGPUサポートに不可欠です。CPUとGPU間のメモリ転送が頻繁に行われる場合 D)、効率が悪くなる可能性があります。過熱 E)により、GPUのパフォーマンスが低下する可能性があります。共有メモリの最適化 B)は役立つ場合もありますが、必ずしもパフォーマンス低下の主な原因とは限りません。

質問: 64

例えば、2つのポートを接続する必要があり、片方がSFP、もう片方がQSFPの場合、25GbE ホストチャネルアダプタを100GbEと25GbEの両方に対応したQSFPポートに接続する場合、どのソリューションがこの要件を最もよく満たすでしょうか？

- A. QSAアダプター。
- B. SFPコネクタ。
- C. SFP-to-1G BASE-T RJ45アダプタ。
- D. 標準QSFP-to-QSFP DACケーブル。

正解: ([正解を表示します](#))

SFP ベースのデバイスを、必要な速度モードをサポートする QSFP ポートに接続する必要がある場合は、QSA アダプタが適切なソリューションです。QSA は QSFP-to-SFP アダプタの略です。スイッチ ポートが 25 GbE などの低いターゲット速度での動作をサポートしている場合、SFP または SFP28 トランシーバまたはケーブルを QSFP または QSFP28 ケージに挿入できます。これは、25 GbE ホスト チャネル アダプタを、物理的には QSFP だが電氣的には 25 GbE に対応しているスイッチ ポートに接続する必要がある混合速度の NVIDIA ネットワーク環境で役立ちます。

25 GbE 動作。SFP コネクタだけでは、QSFP ケージに直接は収まらないため、機械的な不一致は解決しません。SFP-to-1G BASE-T アダプタは 1 GbE 銅線 RJ45 接続用であり、25 GbE の要件を満たしていません。QSFP-to-QSFP DAC ケーブルも、ホスト側が SFP ベースであるため、不適合です。物理層管理中は、クラスタ起動時のリンクダウン状態を回避するために、エンジニアはポート機能、サポートされているケーブルタイプ、トランシーバの互換性、リンク速度、ファームウェアのサポートを確認する必要があります。

質問: 65

サーバーに新しいNVIDIA GPUを物理的にインストールした後、システムを起動します。すると、オペレーティングシステムがGPUを認識しないことがわかりました。カードが正しく装着され、電源が供給されていることは確認済みです。最も可能性の高い原因と解決策は何ですか？ 2つ選択してください)

- A. GPUドライバーが正しくインストールされていないか、ドライバーが全くインストールされていません。解決策 :NVIDIAのウェブサイトから最新のドライバーをダウンロードしてインストールしてください。
- B. マザーボードのBIOS/UEFIがGPUをサポートしていません。解決策 :マザーボードのBIOS/UEFIを最新バージョンにアップデートしてください。
- C. PCIeスロットに不具合があります。解決策 :GPUを別のPCIeスロットに取り付けてみてください。
- D. GPUがオペレーティングシステムと互換性がありません。解決策 :オペレーティングシステムを再インストールしてください。
- E. GPUに不具合があります。解決策 :GPUを製造元に返送してください。

正解: ([正解を表示します](#))

GPUが認識されない最も一般的な原因は、ドライバの欠落または誤り、およびカードをサポートしていない古いBIOS/UEFIです。PCIeスロットの故障も考えられますが、最初のトラブルシューティング手順としては可能性は低いでしょう。ドライバの問題でOSを再イン

ストールする必要はほとんどありません。GPU自体の故障も考えられますが、他の選択肢をすべて試した後に検討すべきです。

質問: 66

During a 72-hour HPL burn-in test on a DGX H100 cluster, one node shows a 15% performance drop after 48 hours. What are the two most likely causes and diagnostic steps?

Pick the 2 correct responses below.

- A. MPI configuration error; rerun with --cpu-affinity adjustments.
- B. Network packet loss; analyze ibdiagnet reports.
- C. Thermal throttling due to cooling issues; check nvidia-smi dmon.
- D. Memory corruption; reboot the node and reduce problem size N.

正解: **B,C** ([コメントを发表する](#))

The two most likely causes are network packet loss and thermal throttling. A performance drop after 48 hours of HPL burn-in is less likely to be a simple launch-time MPI configuration issue, because MPI affinity errors usually appear from the beginning of the run as consistently poor performance. A delayed degradation suggests the system changed state under sustained load. Thermal throttling is a common cause: after many hours, rack cooling imbalance, blocked airflow, high inlet temperature, or fan behavior can cause GPU clocks to drop. nvidia-smi dmon helps monitor GPU temperature, power, utilization, and clocks over time. Network packet loss is also likely in multi-node HPL because HPL depends on heavy communication across the InfiniBand fabric. Link errors, symbol errors, retransmissions, degraded cables, or congestion can reduce sustained performance. ibdiagnet is the correct fabric-level diagnostic tool to collect and analyze InfiniBand health, topology, counters, and link issues. Rebooting and reducing matrix size would hide the symptom rather than diagnose it. Correct burn-in practice is to preserve evidence, inspect thermal telemetry, review network diagnostics, and compare the affected node against healthy peers.

質問: 67

高密度サーバー環境にNVIDIA AIOO GPUのクラスタをデプロイする任務を負っています。サーバーシャーシは電力予算と冷却能力が限られています。ピークパフォーマンス時に電力と冷却インフラストラクチャがGPUワークロードを適切にサポートし、サーマルスロットリングとシステム不安定性のリスクを最小限に抑えられることを検証するために、次のうちの戦略が最も効果的ですか？

- A. GPUメーカーが公表している熱設計電力 (TDP) 仕様のみを依拠し、これらの値に基づいて電力を割り当てます。
- B. 継続的な計算負荷の高いワークロード中に `nvidia-smi` を使用してGPUの温度を監視し、GPUの熱しきい値と比較します。温度がしきい値を下回っている場合は、冷却は適切です。

- C. 想定されるAIワークロードを模倣したストレステスト中に、サーバーの実際の消費電力を測定するために、電力監視ツール（例IPMI、Redfish）を使用します。この値を電源ユニット（PSU）の定格と冷却システムの容量と照合します。
- D. 合成ベンチマーク（例Linpack）を使用してAIワークロードをシミュレートし、ベンチマークのパフォーマンス指標に基づいて消費電力を推定します。
- E. ワークロード実行中のGPUクロック速度を監視します。クロック速度が定格最大速度であれば、電力と冷却は十分です。

正解: **C** ([コメントを发表する](#))

オプションCは最も包括的なアプローチを提供します。TDPは理論上の最大値であり、実際の消費電力を反映するものではありません。温度監視は重要ですが、総消費電力は考慮されません。合成ベンチマークは、AIワークロードを正確に表していない可能性があります。実際の消費電力を監視し、それを電源ユニットの定格および冷却能力と比較することで、最も正確な検証が可能になります。

質問: **68**

You are developing a CI/CD pipeline for deploying A1 models to a Kubernetes cluster with NVIDIA GPUs. Your pipeline includes steps to build, test, and deploy the model as a containerized application. What is the recommended approach for ensuring that the application container image includes the correct NVIDIA drivers and libraries?

- A. Manually installing the NVIDIA drivers and libraries inside the container image during the build process.
- B. Using a base image provided by NVIDIA that includes the necessary drivers and libraries.
- C. Relying on the NVIDIA Container Toolkit to automatically install the drivers and libraries at runtime.
- D. Downloading the latest drivers from NVIDIA's website and adding them to the container image.
- E. Configuring the cluster to download and configure the needed drivers just before deploying your application.

正解: **B** ([コメントを发表する](#))

Using a base image provided by NVIDIA (such as those from NGC) is the recommended approach because it ensures that the container image includes the correct and compatible drivers and libraries, simplifying the build process and reducing potential compatibility issues. The NVIDIA Container Toolkit requires host drivers and handles the injection of the library into the container and not the building of the image.

質問: **69**

An administrator needs to perform a comprehensive pre-production stress test on a DGX H100 system. Which command validates GPU, CPU, memory, and storage components while following NVIDIA's recommended procedure?

- A. `nvidia-smi -q | grep " GPU Stress Test "`
- B. `sudo nvsm stress-test --force`
- C. `stress --cpu $(nproc) --io $(nproc) --timeout 600`
- D. `./gpu_burn 60`

正解: [\(正解を表示します\)](#)

The correct command is `sudo nvsm stress-test --force`. NVIDIA recommends using NVSM for the DGX pre-flight stress test before putting a DGX H100 system into production or after servicing. The documented NVSM stress test can run checks across supported components, including GPUs, CPU, memory, and storage, and the recommended command for all supported components is `sudo nvsm stress-test --force`. `nvidia-smi -q` provides detailed GPU information, but it does not execute a full platform stress test. The Linux `stress` command can load CPU and I/O subsystems, but it is generic and does not validate the DGX platform using NVIDIA's health model. `gpu_burn` may stress GPUs, but it does not cover CPU, system memory, storage, and DGX-specific platform checks in the recommended way. During server bring-up, NVSM is preferred because it understands DGX hardware components and can identify platform health issues before the node is released to production workloads.

質問: 70

A system administrator needs to configure a BlueField DPU and enable RShim on the baseboard management controller (BMC). Which command should be executed?

- A. `ipmitool raw 0x32 0x6a 1`
- B. `systemctl restart rshim`
- C. `systemctl enable bmc-rshim.service`
- D. `scp <path_to_bfb> root@<bmc_ip>:/dev/rshim0/boot`

正解: [A \(コメントを发表する\)](#)

NVIDIA BlueField DPUアーキテクチャでは、RShim (Remote-Shim) インターフェイスが DPU とホストまたは BMC 間の重要な通信チャネルを提供し、通常は初期段階のプロビジョニング、コンソールアクセス、ファームウェアのロードに使用されます。DPU は通常ホストの PCIe バスを介して管理されますが、特定のデータセンター構成では、サーバーのベースボード管理コントローラ (BMC) を介して DPU を帯域外で管理する必要があります。この機能を有効にするには、内部 USB-to-BMC ブリッジを介して RShim 機能を切り替えるための特定の低レベルコマンドを BMC に送信する必要があります。コマンド `ipmitool raw 0x32 0x6a 1` は、NVIDIA DGX および認定システムで BMC-to-DPU RShim パスを有効にするために使用される検証済みの生 IPMI 16 進コードです。有効化されると、BMC は DPU を USB デバイスとして認識し、ホスト CPU の電源がオフまたは応答しない場合でも、管理者は BlueField Boot (BFB) イメージを `/dev/rshim0/boot` にプッシュして OS をインストールできます。オプション B と C は、ドライバが既にロードされ、ハードウェアパスがアクティブになっていることを前提としたホスト側のサービス コマンドですが、ハードウェアパス自体を有効にするには、生の IPMI コマンドが必要です。

質問: 71

BlueField DPIJをインラインセキュリティアプライアンスとして展開する場合、サーバーを通過するすべてのトラフィックを検査する際に、最も効率的でレイテンシが最小となるネットワーク構成はどれですか？

- A. BlueFieldを従来のブリッジとして構成し、2つの物理インターフェース間でトラフィックを転送します。
- B. SR-IOVを使用してホストとDPIJの両方の仮想機能 (VF)を作成し、それらの間でトラフィックをルーティングします。
- C. BlueFieldをルーターとして設定し、サブネット間でレイヤー3転送を実行します。
- D. トラフィックを DPIJ に送信し、検査後に同じ物理インターフェースにループバックするヘアピン構成を実装します。
- E. DPIJがスイッチASICを直接プログラムしてトラフィックを制御するスイッチデバイスベースの構成を利用します。

正解: [E \(コメントを发表する\)](#)

switchdevベースの構成は、DPIJがスイッチASICを直接プログラムするため、最も低いレイテンシを実現します。これにより、従来のブリッジングやルーティングのオーバーヘッドなしに、ハードウェアアクセラレーションによるトラフィックのステアリングと検査が可能になります。ヘアピンニングは不要なレイテンシを発生させ、SR-IOVとブリッジングはソフトウェア処理のオーバーヘッドを増加させます。

質問: 72

GPUDirect Storageを使用してGPUとNVMeドライブ間のメモリへの直接アクセスを有効にしているシナリオを考えてみましょう。GPUDirect Storageを有効にしているにもかかわらず、期待されるパフォーマンス向上が見られないことに気づきました。GPUDirect Storageの最適なパフォーマンスを確保するために、考えられる原因と確認すべき構成は何ですか？該当するものをすべて選択してください。

- A. NVMeドライブがRAID 0構成で正しく構成されていることを確認してください。
- B. NVMeドライブがPCIe Gen4またはGen5経由でシステムに接続されていることを確認してください。
- C. CUDAドライバのバージョンがGPIJDirect Storageと互換性があることを確認してください。
- D. ファイルシステムが直接入出力（例「directio」マウントオプションの使用）をサポートしているかどうかを確認します。
- E. Disable CPU-side caching to force all I/O operations to go directly to the GPU memory.

正解: [\(正解を表示します\)](#)

Explanation:GPUDirect Storage requires PCIe Gen4/Gen5 for sufficient bandwidth (B). The CUDA driver must be compatible with GPUDirect Storage (C). Direct I/O support in the file system is essential to bypass the OS cache and allow direct GPU access (D). RAID 0 (A) is about storage speed but not directly related to GDS functionality. Disabling CPU-side

caching (E) is usually detrimental as it can reduce overall system performance. Note, this is not always bad but needs to be tested depending on application.

質問: 73

Consider a scenario where you are setting up a high-performance computing cluster with several GPU-accelerated nodes using Slurm as the resource manager. You want to ensure that jobs requesting GPUs are only scheduled on nodes with the appropriate NVIDIA drivers and CUDA toolkit installed. How can you achieve this within Slurm?

A. Use Slurm's 'GresTypeS configuration option in 'slurm.conf to define a generic resource type called 'gpu' and then configure each node to advertise the available GPUs. Slurm will automatically ensure that jobs requesting GPUs are only scheduled on nodes with the 'gpu' resource.

B. Create a custom Slurm script that checks for the presence of the NVIDIA driver and CUDA toolkit before submitting a job to a node. If the requirements are not met, the job is rejected.

C. Use Slurm's node features to tag nodes with the "Feature=" keyword in 'slurm.conf. For example, tag nodes with GPUs as 'Feature=gpu' . Jobs can then request nodes with the 'gpu' feature using the option.

D. Install the NVIDIA Data Center GPU Manager (DCGM) on each node and configure Slurm to query DCGM for GPU availability and health. Slurm will then only schedule jobs on healthy and available GPUs.

E. Utilize Slurm's Prolog and Epilog scripts to dynamically install the necessary NVIDIA drivers and CUDA toolkit on each node before and after a job runs. This ensures that the required software is always available.

正解: ([正解を表示します](#))

Using Slurm's node features is the most straightforward and recommended approach for tagging nodes with specific capabilities. The '-constraint' option allows jobs to request nodes with particular features. GresTypeS can be used, but node features provide more flexibility and control. Installing drivers dynamically is impractical and inefficient. DCGM is primarily for monitoring, not core scheduling requirements.

質問: 74

複数のNVIDIA A100 GPUを搭載したAMD EPYCサーバーのクラスター上で、大規模な分散トレーニングジョブを実行しています。ジョブスケジューリングにはSlurmを使用しています。トレーニングプロセスは、ネットワーク接続に関連するNCCLエラーで頻繁に失敗します。この環境でNCCLのネットワーク通信の信頼性を向上させるには、どのような対策を講じればよいのでしょうか？最も適切な回答を選択してください。

A. InfiniBandまたはRoCEネットワークが正しく構成され、すべてのサーバーがネットワーク経由で相互に通信できることを確認してください。NCCL構成でネットワークインターフェース名とIPアドレスを確認してください。

- B. Slurmの `srun` コマンドに `mpi=pmi2` オプションを付けてトレーニングジョブを起動します。これにより、SlurmがMPI環境を適切に初期化し、NCCL環境変数を設定することが保証されます。
- C. 'NCCL CONNECT TIMEOUT' および *NCCL TIMEOUT 環境変数を増やして、より長いネットワーク遅延を許容します。
- D. すべてのサーバーのファイアウォールを無効にして、制限のないネットワーク通信を許可します。
- E. バッチサイズを小さくして、ネットワーク経由で転送されるデータ量を削減します。

正解: [\(正解を表示します\)](#)

ネットワーク構成が正しいことを確認することが最も重要なステップです。 `srun` に `mpi=pmi2` オプションを付けると、Slurmによって設定されたNCCL環境変数が自動的に処理され、適切な接続が確保されます。タイムアウト時間を長くすると、一時的なネットワークの問題が障害を引き起こすことなく解決されます。ファイアウォールを無効にするとセキュリティリスクが発生します。バッチサイズを小さくするとデータ量は減りますが、コアとなるネットワーク接続の問題は解決されません。

質問: 75

あなたは、BlueField DPIJを活用してネットワークトラフィックに対してディープパケットインスペクション (DPI) を実行するセキュリティアプリケーションをデプロイしようとしています。

Your application requires access to the raw packet data, including the Ethernet headers. Which of the following programming models or APIs is most suitable for accessing raw packet data on the BlueField DPU with minimal overhead?

- A. Using standard Linux socket APIs (e.g., 'socket()', 'recvfrom()') in user space.
- B. using DPDK (Data Plane Development Kit) with the BlueField DPU's PMD (Poll Mode Driver).
- C. Using Netfilter hooks in the Linux kernel to intercept packets.
- D. Using libpcap to capture packets from the network interface.
- E. Using the 'tcpdump' command and parsing its output.

正解: **B** ([コメントを发表する](#))

DPDK provides a framework for high-performance packet processing in user space, bypassing the kernel's networking stack. The BlueField DPU's PMD (Poll Mode Driver) allows DPDK applications to directly access the network interface with minimal overhead. This is ideal for DPI applications that need to process a large volume of packets at high speed. Standard socket APIs and Netfilter hooks involve kernel intervention, which can introduce significant overhead. Libpcap is a packet capture library, not designed for high-performance packet processing. 'tcpdump' is a command-line tool, not a programming API.

質問: 76

故障したNVIDIA GPUを交換した後、システムは起動し、`nvidia-smi`コマンドで新しいカードが検出されます。しかし、CUDAプログラムを実行すると、「CUDA対応デバイスが検出されません」というエラーで失敗します。正しいドライバがインストールされ、GPUが正しく装着されていることは確認済みです。この問題の最も可能性の高い原因は何でしょうか？

- A. 新しいGPUは既存のシステムBIOSと互換性がありません。
- B. CUDAツールキットが新しいGPUを使用するように正しく構成されていません。
- C. 'LD_LIBRARY_PATH' 環境変数が正しく設定されていません。
- D. CUDAプログラムを実行しているユーザーには、GPUにアクセスするために必要な権限がありません。
- E. ACPI 設定が欠落しているか間違っているため、GPIO がシステムによって正しく初期化されませんでした。

正解: (正解を表示します)

`nvidia-smi`でGPIOが認識されているにもかかわらず「CUDA対応デバイスが検出されません」というエラーが発生する場合は、CUDAがカードにアクセスできない低レベルのシステム問題が原因です。このような場合、ACPI (Advanced Configuration and Power Interface) の設定ミスが原因であることがよくあります。ACPIはデバイスの初期化と電源管理を処理します。ACPIが新しいGPUを適切に構成しない場合、CUDAプログラムはGPUにアクセスできません。まず最初に行うべきことは、ACPI構成を確認して修正することです。これには、システムBIOS/IJEFIでPCIデバイス、特にGPU/アクセラレータに関連する設定が適切であることを確認することが含まれます。LD_LIBRARY_PATHはCUDAライブラリのランタイムリンクに影響しますが、基本デバイスの検出には影響しません。`nvidia-smi`が動作していることから、ユーザー権限が原因である可能性は低いです。

有効的な**NCP-AII**問題集はJPNTTest.com提供され、**NCP-AII**試験に合格することに役に立ちます！JPNTTest.comは今最新**NCP-AII**試験問題集を提供します。JPNTTest.com NCP-AII試験問題集はもう更新されました。ここで**NCP-AII**問題集のテストエンジンを手に入れます。最新版のアクセス、<https://www.jpntest.com/shiken/NCP-AII-mondaishu>
125問、30%ディスカウント、特別な割引コード: **JPNshiken**」

質問: 77

データサイエンティストから、DGX A100サーバーでのトレーニングパフォーマンスが過去1週間で著しく低下したとの報告がありました。`nvidia-smi`ではすべてのGPUが正常に動作していることが示されていますが、`nvprof`では `cudaMemcpy`の実行時間が大幅に増加していることが明らかになりました。最も可能性の高いボトルネックは何でしょうか？

- A. CPUに高い負荷がかかっており、システムメモリの帯域幅をめぐる競合が発生しています。
- B. PCIeバスが飽和状態になり、CPUとGPU間のデータ転送速度が制限されます。
- C. GPU間のNVLink接続が失敗し、PCIe経由でのデータ転送が強制されます。
- D. GPUが過熱し、サーマルスロットリングが発生してメモリ転送速度が低下しています。
- E. ストレージシステムが遅く、データ読み込みと前処理が遅延しています。

正解: [\(正解を表示します\)](#)

ScudaMemcpyの実行時間が長くなっている場合、CPUとGPU間のデータ転送、またはGPUメモリ自体にボトルネックが発生していることを示しています。PCIeの飽和やNVLink接続の不具合も原因となる可能性があります。多くの場合、CPUの負荷が高いことが主な原因です。CPU負荷の高い前処理タスクや一般的なシステム負荷によってシステムメモリの競合が発生し、GPUへのデータ転送速度が低下することがあります。NVLinkが原因の可能性は低いと考えられます。なぜなら、nvidia-smiコマンドを実行すれば、GPU間でエラーや速度低下が報告されるはずだからです。

質問: 78

管理者は、BlueFieldイメージをターゲットDPUに手動で展開する必要があります。管理者は新しいイメージファイルをダウンロードし、それをハードウェアに書き込む必要があります。管理者はどのコマンドを使用すべきでしょうか？

- A. /opt/mellanox/mlnx-fw-updater/mlnx_fw_updater.pl
- B. apt install doca-runtime
- C. dd if=/root/bf.image of=/dev/bf bs=4096k
- D. bfb-install --rshim

正解: [D \(コメントを发表する\)](#)

正しいコマンドは bfb-install --rshim で、通常は BFB イメージパスと適切な RShim デバイスとともに使用します。例: `sudo bfb-install --rshim rshim0 --bfb < image_path.bfb >`。BlueField ソフトウェア イメージは一般的に BFB ファイルとして展開され、RShim はイメージを BlueField デバイスにプッシュするために使用されるホスト側のパスを提供します。NVIDIA のドキュメントには、bfb-install ユーティリティが RShim パッケージに含まれており、インストールの進行状況を報告しながら BFB イメージを BlueField 側にプッシュするために使用されると記載されています。mlnx_fw_updater.pl ツールはファームウェアの更新用であり、完全な BlueField OS イメージの展開用ではありません。apt install doca-runtime は DOCA ランタイム パッケージをインストールしますが、BlueField イメージをフラッシュしません。dd を直接架空のデバイスパスに使用することは安全ではなく、BlueField ブート ストリーム イメージを展開するためのサポートされた方法ではありません。起動時にサポートされている BFB インストール ワークフローを使用すると、DPU が有効な署名付きイメージを起動し、既知の動作状態に入ることが保証されます。

質問: 79

NVIDIA GPU上でCUDAアプリケーションを実行している状況を考えてみましょう。アプリケーションは正常にコンパイルされますが、実行時に「CUDA ERROR ILLEGAL ADDRESS」エラーが発生してクラッシュします。コードを注意深く確認しましたが、明らかなメモリ範囲外アクセスは見つかりません。このエラーの原因を特定するには、どのような高度なデバッグ手法が有効でしょうか？

- A. 実行時にメモリアクセスエラーを検出するには、「cuda-memcheck」を使用します。
- B. CUDAデバッガ（cuda-gdb）を使用してコードをステップ実行し、変数値とメモリ内容を調べます。
- C. NVIDIA Nsight Systems を利用してアプリケーションをプロファイリングし、メモリ割り当てパターンを特定します。
- D. GPU 上で ECC (エラー訂正コード) メモリを有効にして、メモリのエラーを検出して修正します。
- E. CUDAカーネルで使用されるブロックサイズを小さくして、共有メモリの競合が発生する可能性を低減します。

正解: [\(正解を表示します\)](#)

「cuda-memcheck」 A) は、メモリアクセスエラーを検出するために特別に設計されています。「cuda-gdb」 B) は、詳細なコード検査を可能にします。NVIDIA Nsight Systems C) は、エラーの原因となるメモリ割り当ての問題を特定するのに役立つプロファイリング情報を提供します。ECC D) を有効にすると症状は隠蔽される可能性があります、根本的なコードエラーは修正されません。また、ECCはデフォルトで有効になっているため、これは的外れです。ブロックサイズを小さくする E) ことは、共有メモリが関係しており、オーバーフローする可能性がある場合にのみ有効です。

質問: 80

You are tasked with optimizing an Intel Xeon scalable processor-based server running a TensorFlow model with multiple NVIDIA GPUs.

You observe that the CPU utilization is low, but the GPU utilization is also not optimal. The profiler shows significant time spent in 'tf.data' operations. Which of the following actions would MOST likely improve performance?

- A. Increase the number of threads used for CPU-bound operations in TensorFlow using 'tf.config.threading.set_intra_op_parallelism_threads()'.
- B. Enable XLA (Accelerated Linear Algebra) compilation in TensorFlow.
- C. Use 'tf.data.AUTOTUNE' to allow TensorFlow to dynamically optimize the data pipeline.
- D. Reduce the global batch size to improve memory utilization.
- E. Upgrade the server's network adapter to a faster interface, such as 100Gb

正解: [\(正解を表示します\)](#)

'tf.data' performance issues often stem from inefficient data pipelines. 'tf.data.AUTOTUNE' allows TensorFlow to dynamically optimize the pipeline by adjusting parameters such as prefetch buffer size and the number of parallel calls to transformation functions. XLA

compilation optimizes graph execution, but 'tf.data' issues need to be addressed first. Increasing CPU threads might help but 'AUTOTUNE' is more specific to the problem. A smaller batch size could negatively impact GPU utilization. Network upgrades are irrelevant as the problem lies within the server.

質問: 81

エンジニアは、BlueField-3 DPUのBMC上のすべてのコンポーネント

(ATF、BSP、NIC、UEFI)の現在のファームウェアバージョンを確認する必要があります。この情報を提供するRedfish APIコマンドはどれですか？

A. `mlxconfig -d <dev> q`

B. `curl -k -u root:<password> -X GET https://<DPU-BMC-IP>/redfish/v1/UpdateService/FirmwareList`

C. `mstflint -d <PCI_ID>` クエリ完全版

D. `curl -k -u root:<password> -X GET https://<DPU-BMC-IP>/redfish/v1/UpdateService/ファームウェアインベントリ`

正解: **D** ([コメントを发表する](#))

最新のNVIDIA BlueField DPUには、業界標準のRedfish APIによる帯域外管理をサポートする統合ベースボード管理コントローラ (BMC)が搭載されています。ホストOSから `mlxconfig` (オプションA) や `mstflint` (オプションC) などのCLIツールを使用してNICファームウェアをチェックすることはできますが、ARM Trusted Firmware (ATF)、ボードサポートパッケージ (BSP)、DPUのUEFIブートローダーといったBMC固有のコンポーネントを簡単に照会することはできません。Redfish規格では、ハードウェアインベントリ用の共通URIが指定されています。FirmwareInventoryエンドポイント (オプションD) は、DPU上のファームウェアで制御可能なすべてのコンポーネントのバージョン情報を含む包括的なJSONオブジェクトを取得するための正しいRESTfulパスです。これは、自動化されたデータセンター管理システム (NVIDIA Base Command Managerなど) が、ステージングフェーズ中にDPUが正しい「ロールデンイメージ」バージョンであることを確認するための推奨される方法です。

FirmwareList」 (オプションB) は、この特定のデータに対する標準的なRedfish URIではないことに注意してください。

質問: 82

BlueField-2 DPUによって高速化されたNVMe-oFトラフィックのパフォーマンス問題のトラブルシューティングを行っています。RDMA構成に問題があると疑っています。パケット損失や輻輳などの潜在的なRDMAの問題を診断するために、次の `perfquery` コマンドのうちどれが最も関連性の高い情報を提供しますか？

A. `'perfquery -x'` (一般的なリンク統計情報)

B. `'perfquery -s:'` (スイッチ統計)

C. `'perfquery'` (QoS統計情報)

D. `'perfquery -P'` (パケット損失と輻輳を含むポートカウンタ)

E. 'perfquery -G' (グローバルカウンター)

正解: [\(正解を表示します\)](#)

「perfquery -P」コマンドは、ポートレベルでのパケット損失、輻輳、その他のRDMA関連メトリックに関する重要な情報を含むポートカウンタを提供します。これは、NVMe-oF環境におけるRDMA関連のパフォーマンス問題を診断する上で最も適切なコマンドです。その他のオプションでは、より限定的または関連性の低い情報しか得られません。

質問: 83

あなたは、異なる物理サーバー上で動作する複数のDockerコンテナにまたがる複数のGPUを使用する分散型深層学習アプリケーションを開発しています。各コンテナがそれぞれのホスト上のGPUにアクセスして利用できるようにするにはどうすればよいでしょうか？

- A. すべてのホストマシンに同じバージョンのNVIDIAドライバーをインストールし、コンテナ間でCUDAライブラリを共有するためのネットワークファイルシステム (NFS)を設定します。
- B. 各ホストマシンにNVIDIA Container Toolkitがインストールされ、構成されていることを確認し、Kubernetesなどのコンテナオーケストレーションプラットフォームを使用して、デプロイとGPU割り当てを管理します。
- C. 各コンテナを手動で設定し、CUDA_VISIBLE_DEVICES」環境変数を使用して、それぞれのホストで使用するGPUを指定します。
- D. Docker Swarmを使用し、「docker-compose.yml」ファイルでGPUリソースの制約を指定して、各サービスにGPUを割り当てます。
- E. カスタムDockerネットワークを作成し、各コンテナがネットワークのゲートウェイをデフォルトのGPUデバイスとして使用するよう設定します。

正解: [\(正解を表示します\)](#)

分散型GPU利用のための最も堅牢なソリューションは、Kubernetes (B)のようなコンテナオーケストレーションプラットフォームとNVIDIA Container Toolkitを組み合わせることです。Kubernetesは、複数のノード間でのスケジューリング、リソース割り当て (GPUを含む)、およびネットワーク処理を行います。

NVIDIA Container Toolkit は、各コンテナがホスト上の GPU にアクセスできるようにします。C)は便利ですが、マルチサーバー展開には十分ではありません。Docker Swarm (D)は機能しますが、Kubernetes のような高度な GPU スケジューリング機能がありません。NFS 共有 (A)は不要であり、パフォーマンスのボトルネックを引き起こす可能性があります。カスタム Docker ネットワーク (E)は、GPU アクセスに直接対応していません。

質問: 84

BlueField-2 DPUを搭載したサーバー群を管理しています。1台のサーバーで断続的なネットワーク接続の問題が発生しています。DPUのファームウェアに問題があるのではないかと疑っています。BlueField-2 DPUの現在のファームウェアバージョンを確認する最も信頼性の高い方法はどれですか？

- A. 起動中にファームウェア関連のメッセージがないかシステムログを確認してください。
- B. デバイスの状態とファームウェアのバージョンを照会するには、`fnst status` コマンドを使用します。
- C. `/proc/driver/mlx4_core/versiorf` ファイルを調べます。
- D. IPMIまたはRedfishを介してDPIJのBMC (ベースボード管理コントローラ)に問い合わせます。
- E. DPIJに関連付けられたネットワークインターフェースで `ethtool -i` を実行します。

正解: ([正解を表示します](#))

`fnst status` コマンドが最も信頼性の高い方法です。このコマンドはMellanox OFEDツールの一部であり、DPLJを含むMellanoxデバイスの状態を照会するために特別に設計されており、ファームウェアバージョンを正確に報告します。他の方法でも手がかりが得られる可能性はありますが、`fnst status` コマンドが決定的な情報源となります。

質問: 85

NCCLを用いた分散型深層学習トレーニングに使用されるマルチGPU Intel Xeonサーバーにおいて、GPU間の通信パフォーマンスを向上させるのに効果的な手法は次のうちどれですか？

- A. GPU間でのPCIeピアツーピア転送を有効にする。
- B. 利用可能な場合は、InfiniBandまたはRoCEインターコネクトを利用する。
- C. システムRAMのサイズを増やして、ディスクへのデータ転送を最小限に抑えます。
- D. NCCL が正しいネットワーク インターフェイスとトランスポート プロトコル (例: 1B、Socket) を使用するように設定します。
- E. 一貫したパフォーマンスを維持するために、CPU周波数スケーリングを無効にします。

正解: ([正解を表示します](#))

GPU間の通信を改善するには、GPU間でデータを転送するために使用されるネットワークを最適化する必要があります。PCIeピアツーピア、InfiniBand/RoCE、および適切なNCCL構成はすべて、通信速度の向上に貢献します。RAMサイズを増やすことはデータキャッシュに役立ちますが、GPU間の通信速度に直接影響するわけではありません。CPU周波数スケーリングを無効にするのは、CPUのパフォーマンス安定性に関するものであり、GPU間の通信に直接影響するものではありません。

質問: 86

ClusterKitが「GPUホストのレイテンシがしきい値を超えました」と報告した場合、ハードウェア障害を特定するためにどのNVIDIA診断ツールを使用すべきですか？

- A. ClusterKitを`--stress=gpu 60`オプション付きで再実行してテスト時間を延長してください。
- B. `nvidia-smi topo -m` コマンドで GPU トポロジーの接続を検査します。
- C. DCGM診断 `dcgmi diag -r 2`
- D. `ib_write_bw` はノード間のInfiniBand帯域幅を測定します

正解: **B** ([コメントを发表する](#))

NVIDIA DGX または HGX システムにおける「GPU ホスト遅延」の問題は、多くの場合、PCIe アフィニティの誤り、または NUMA (非均一メモリ アクセス) マッピングの最適化不足が原因です。GPU がローカル PCIe スイッチ/ルート コンプレックス上にない CPU コアまたは HCA と通信するように強制されると、データが QPI/UPI プロセッサ間リンクを通過する必要があるため、遅延が大幅に増加します。コマンド `nvidia-smi topo -m` は、システムの内部トポロジの詳細なマトリックスを提供し、GPU、CPU、および NIC がどのように接続されているかを示します。接続が単一の PCIe スイッチ (PIX)、複数のスイッチ (PXB)、または CPU (SYS) を介しているかどうかを識別します。

このマップを検査することで、管理者はソフトウェアプロセスが誤った NUMA ノードに固定されているか、ハードウェアパスが予期せず劣化しているかを特定できます。DCGM オプション C) はコンポーネントの状態を確認するのに適していますが、特定のレイテンシ「しきい値」警告の原因となる論理アフィニティパスと物理アフィニティパスのマッピングは行いません。

質問: **87**

あなたは、新しい NVIDIA Quantum-2 InfiniBand スイッチの導入において、トランシーバーを選定する任務を負っています。主な要件は、短距離 (最大 50 メートル) で 400 Gbps の帯域幅を維持しながら、消費電力を最小限に抑えることです。このシナリオにおいて、最も電力効率に優れたトランシーバーの種類はどれでしょうか？

- A. QSFP-DD LR8
- B. QSFP-DD DR4
- C. QSFP-DD SR8
- D. QSFP-DD AOC
- E. QSFP-DD SR4

正解: ([正解を表示します](#))

SR4 トランシーバーは、他の 400 GbE トランシーバーに比べて消費電力が比較的低いことで知られています。これは、よりシンプルな変調方式と短い伝送距離を採用しているため、信号増幅と処理に必要な電力が少なく済むからです。LR8 と DR4 はより長距離向けに設計されており、消費電力も高くなります。AOC は便利ですが、一般的には最も電力効率の良い選択肢ではありません。SR8 は SR4 よりも消費電力が若干高い場合がありますが、特定の状況ではより優れたパフォーマンスを発揮します。

質問: **88**

NVIDIA A100 GPU を、以前に MIG モードを設定した後、非 MIG モードの状態にリセットしたい場合を考えてみましょう。以下のどの手順が必要ですか？

- A. `'nvidia-smi -set-mig-mode=disable -i 0'` を実行してからシステムを再起動します。
- B. `'nvidia-smi -destroy-mig-config -i 0'` を実行し、次に `'nvidia-smi -set-mig-mode=disable -i 0'` を実行し、最後に再起動します。

C. 'nvidia-smi -set-mig-mode=disable -i 0' を実行してから 'nvidia-smi -reset-default-mig-mode -i' を実行します。

D. 'nvidia-smi -set-mig-mode=disable -i' を実行してから、システムをシャットダウンし、GPU を物理的に取り外して再取り付けします。

E. 'nvidia-smi -set-mig-mode=disable -i 0' を実行し、次に 'nvidia-smi -i 0 -migrr 0' を実行し、最後に再起動します。

正解: ([正解を表示します](#))

GPU を非 MIG モードにリセットするには、コマンド `nvidia-smi -set-mig-mode=disable -i 0` を実行し、その後システムを再起動する必要があります。これにより、次回の起動プロセス中に変更が適用されます。MIG モードを無効にするだけであれば、MIG 設定の削除や GPU の物理的な再インストールは不要です。

質問: 89

BCMクラスタのセットアップ中に、エンジニアは高可用性を確保するために、DGXノード上でボンディングされたネットワークインターフェイスを構成する必要があります。2つの物理NICを持つbond0インターフェイスを適切に構成するcmlshコマンドシーケンスはどれですか？

A. デバイス使用dgx001; インターフェイス追加vlan vlan100; 親bond0を設定; モード1を設定; ネットワークinternalnetを設定

B. デバイス使用dgx001; インターフェイス追加bond0; インターフェイスenp225s0f1np1 enp97s0f1np1を追加; モード1を設定; ネットワークinternalnetを設定

C. デバイス使用dgx001; インターフェイス設定enp225s0f1np1 ネットワークinternalnet; インターフェイス設定enp97s0f1np1 ネットワークinternalnet

D. デバイス使用dgx001; インターフェイス削除enp225s0f1np1; インターフェイス削除enp97s0f1np1

正解: ([正解を表示します](#))

NVIDIA Base Command Manager (BCM) では、管理トラフィックとストレージトラフィックに NIC ボンディング (リンクアグリゲーション) による冗長性が必要となることがよくあります。cmlsh ユーティリティは階層的なコマンド構造を使用してノード構成を変更します。ボンディングを作成するには、まず管理者は特定のノードのインターフェイス構成 (device use <node>) に移動する必要があります。正しい手順としては、タイプが `bond` の新しいインターフェイスを追加します (interfaces add bond bond0)。ボンディングオブジェクトが作成されたら、append interfaces コマンドを使用して物理スレーブインターフェイスをボンディングオブジェクトに関連付ける必要があります。最後に、ボンディングモード (アクティブ/バックアップの場合はモード 1、LACP の場合はモード 4) と論理ネットワーク割り当てを定義する必要があります。オプション B はこのロジックに正しく従っています。

ネットワークを物理ポートに直接割り当てる (オプションC) と、統合ボンディングIPが使用できなくなります。また、オプションAでは、基盤となるボンディングが確立される前に

VLANを追加しようとするかもしれませんが、これは誤りです。この構成は、単一の管理ケーブルまたはスイッチポートが故障した場合でも、制御プレーンへのアクセスを維持するために不可欠です。

質問: 90

You are experiencing link flapping (frequent up/down transitions) on several InfiniBand links in your AI infrastructure. This is causing intermittent connectivity issues and performance degradation. What are the MOST likely causes of this issue, and what steps should you take to troubleshoot and resolve it? (Select TWO)

- A. Incorrect MTU (Maximum Transmission Unit) configuration on the affected interfaces.
- B. Faulty or damaged cables, connectors, or transceivers.
- C. Software bugs in the operating system or InfiniBand drivers.
- D. Mismatched link speeds or duplex settings between connected devices.
- E. Excessive broadcast traffic causing congestion.

正解: [\(正解を表示します\)](#)

Link flapping is most commonly caused by physical layer issues (faulty cables, connectors, or transceivers) or configuration mismatches (link speeds or duplex settings).

Troubleshooting should focus on inspecting the physical connections and verifying that the link speed and duplex settings are correctly configured on both ends of the link. While MTU issues and software bugs can cause network problems, they are less likely to directly cause link flapping. Excessive broadcast traffic can cause performance issues but is less likely to result in frequent link up/down transitions.

質問: 91

BlueField-2 DPII を構成してネットワーク仮想化機能をオフロードする作業を任せられました。vPorts を構成する前に、ホストと DPU で SR-IOV が正しく有効化され、機能していることを確認するために最も重要な Mellanox OFED コマンドはどれですか？

- A. リンク速度とFEC設定を確認するには `mnlxlink` を使用します。
- B. Mellanox デバイスとそのファームウェアの状態を確認するには、`mst status` を使用します。
- C. 仮想関数 (VF) の存在とその割り当てを確認するには、`lspci -vv` を使用します。
- D. `ifconfig` または `ip addr show` を使用して、ネットワークインターフェースへの IP アドレスの割り当てを確認します。
- E. `ethtool -i` でドライバ情報を確認します。

正解: **C** ([コメントを发表する](#))

`lspci -vv` は最も重要なコマンドです。このコマンドは仮想機能 (VF) の存在を表示し、SR-IOV が有効になっていることを確認します。また、割り当てられたデバイス ID も表示するため、ホストと DPII がネットワークインターフェースに VF を正しく認識して割り当てていることを検証できます。他のコマンドは一般的なネットワーク診断に役立ちますが、`lspci -vv` のように SR-IOV の機能を直接検証することはできません。

有効的な**NCP-AII**問題集はJPNTTest.com提供され、**NCP-AII**試験に合格することに役に立ちます！JPNTTest.comは今最新**NCP-AII**試験問題集を提供します。JPNTTest.com NCP-AII試験問題集はもう更新されました。ここで**NCP-AII**問題集のテストエンジンを手に入れます。最新版のアクセス、<https://www.jpntest.com/shiken/NCP-AII-mondaishu> **125問、30%ディスカウント**、特別な割引コード: **JPNshiken**」

質問: 92

You observe high latency and low bandwidth between two GPUs connected via an NVLink switch. You suspect a problem with the NVLink link itself. Which of the following methods would be the most effective in diagnosing the physical NVLink link health?

- A. Using 'iperf3' to measure network throughput between the servers.
- B. Running a CUDA-aware memory bandwidth test specifically designed for NVLink.
- C. Examining system logs for NVLink-related error messages.
- D. Using 'ping' to check network connectivity between the servers.
- E. Physically inspecting the NVLink cables for damage.

正解: ([正解を表示します](#))

A CUDA-aware memory bandwidth test can specifically measure the NVLink link's performance. System logs can reveal hardware-level errors. Physical inspection can identify damaged cables. 'iperf3' and 'pings' are network-level tools and don't directly test the NVLink link. Checking for error messages in System Logs also helps identify potential issues related to the NVLink switch and the link connections.

質問: 93

セキュアブートが有効になっているシステムでGPUアクセラレーションコンテナを実行すると、「アクセス拒否」エラーが発生します。この問題を解決するにはどうすればよいでしょうか？

- A. MOKを登録し、NVIDIAカーネルモジュールに署名します。
- B. NVIDIAランタイムなしでDockerを再インストールします。
- C. 不要なセキュリティポリシーを緩和するために、SELinuxを無効にします。
- D. 管理者権限を得るために、sudo を使用して Docker を実行します。

正解: A ([コメントを发表する](#))

UEFIセキュアブートが有効になっているシステムでは、Linuxカーネルはロードされるすべてのカーネルモジュールに対して「署名検証」ポリシーを厳密に適用します。NVIDIAドライバは、ハードウェアとNVIDIAコンテナランタイム間のインターフェースを提供する複数のモジュール (nvidia.koやnvidia-uvmm.koなど) で構成されています。これらのモジュールがシステムのファームウェアによって信頼されているキーで署名されていない場合、カーネルはモジュールのロードをブロックし、コンテナが/dev/nvidia/ctdまたはGPUデバイスノー

ドにアクセスしようとしたときに「Permission Denied」エラーが発生します。ホストのセキュリティを損なうことなくこの問題を解決するには、管理者はマシンオーナーキー (MOK) を使用する必要があります。モジュールはこのキーを使用して署名する必要があります。キー自体はブートプロセス中にMOK管理インターフェースを介してシステムのセキュアブートデータベースに「登録」する必要があります。これにより、ドライバが機能するための「信頼の連鎖」が確立されます。SELinuxを無効にする (オプションC) か、sudoを使用する (オプションD) だけではこの問題は解決しません。なぜなら、ブロックはユーザー空間の権限レベルではなく、カーネルとファームウェアのインターフェースレベルで発生しているからです。

質問: 94

CPUとGPUの両方のリソースを使用するクラウドネイティブアプリケーションを構築しています。GPUを搭載していないノードにCPU負荷の高いタスクを、GPUを搭載したノードにGPU負荷の高いタスクを割り当てることで、リソース利用率とコストを最適化したいと考えています。Kubernetesでこのようなノード選択とワークロード配置を実現するにはどうすればよいでしょうか？

- A. ノードアフィニティルールを使用して、CPU負荷の高いタスクをGPU搭載ノードに、GPU負荷の高いタスクをGPU非搭載ノードにスケジュールします。
- B. ノードアフィニティルールを使用して、CPU負荷の高いタスクをGPUのないノードに、GPU負荷の高いタスクをGPUのあるノードにスケジュールします。
- C. テイントとトレランスを使用して、GPUのないノードをCPU負荷の高いタスクに、GPUのあるノードをGPU負荷の高いタスクに割り当てます。
- D. リソースクォータを使用して、GPUを搭載したノードで使用可能なCPUリソースと、GPUを搭載していないノードで使用可能なGPUリソースを制限します。
- E. ラベルを使用して、CPUおよびGPU負荷の高いノードを識別します。

正解: **B** ([コメントを发表する](#))

ノードアフィニティルールを使用すると、Podをスケジュールできるノードに制約を指定できます。ノードアフィニティルールを使用することで、CPU負荷の高いタスクはGPUのないノードに、GPU負荷の高いタスクはGPUのあるノードにスケジュールされるようになります。これにより、リソースの利用効率とコストが最適化されます。テイントやトレランスも使用できますが、アフィニティの方が柔軟性があります。リソースクォータはリソースの使用量を制限しますが、配置を制御するものではありません。

質問: 95

高度にセキュリティが確保された環境にBlueField OSを導入する場合、OSの導入中および導入後に考慮すべきセキュリティ対策として最も重要なものはどれですか？

- A. セキュアブートを有効にして、ブートプロセス中に信頼できるコードのみが実行されるようにします。

- B. 必要なサービスのみネットワークアクセスを制限するために、強力なファイアウォールを設定します。
- C. BlueField OS およびインストールされているすべてのソフトウェアを定期的に更新し、セキュリティの脆弱性を修正します。
- D. 攻撃対象領域を縮小するために、不要なサービスとポートを無効にする。
- E. すべてのユーザーアカウントに対して厳格なパスワードポリシーを導入します。

正解: ([正解を表示します](#))

挙げられたすべてのオプションは、セキュリティ体制の強化に大きく貢献します。セキュアブートは起動の完全性を保証し、ファイアウォールは不正なネットワークアクセスを制限し、定期的なアップデートは脆弱性を修正し、不要なサービスを無効にすることで攻撃対象領域を最小限に抑え、強力なパスワードポリシーはユーザーアカウントを保護します。これらはすべて、安全な環境において不可欠です。

質問: 96

A1サーバーでGPU負荷が高いときにカーネルパニックが頻繁に発生します。dmesgには次のエラーが表示されます。NVRM: Xid (PCI:0000:3B:00): 79, pid=..., name=..., GPU has fallen off the bus.」この問題の原因として最も可能性が低いのは次のうちどれですか？

- A. GPUへの電力供給が不十分なため、負荷がかかった際に不安定になります。
- B. GPUとマザーボードを接続するPCIeライザーケーブルが緩んでいるか破損している。
- C. NVIDIAドライバーのバグにより、GPUが不安定になります。
- D. GPUを安定動作限界を超えてオーバークロックすること。
- E. CPUに不具合があります。

正解: ([正解を表示します](#))

GPUがバスから外れました」というエラーメッセージは、GPUとマザーボードまたは電源ユニットとの接続にハードウェア関連の問題があることを強く示唆しています。電力不足、ライザーケーブルの緩み、ドライバのバグ、オーバークロックなどが原因となる可能性があります。CPUの故障はシステムの不安定性を引き起こす可能性はありますが、GPUがバスから外れることとは直接的な関連性が低いため、この状況では最も可能性の低い原因と言えます。

質問: 97

システム管理者がサーバーにGPUをインストールしようとしています。デバイスを損傷しないようにする必要があります。どのようなアイテムを使用すべきでしょうか？

- A. 静電気放電防止ストラップ
- B. 手袋
- C. 保護フィルム
- D. 電動ドライバー

正解: A ([コメントを发表する](#))

H100やA100などの高性能NVIDIA GPUは、静電気放電 (ESD)に非常に敏感です。人間が感じることさえできない静電気の火花 (3,000ボルト未満)でも、GPUダイやHBM (高帯域幅メモリ)モジュール内の微細回路に永久的な損傷を与える可能性があります。内部サーバーコンポーネントを扱う技術者にとって、アンチESDストラップ (またはリストストラップ)は必須の安全アイテムです。これは、技術者を接地することで、ハードウェアに触れる前に体に蓄積された静電気を安全に放電させる仕組みです。手袋 (オプションB)は鋭利なエッジから保護してくれるかもしれませんが、ESD対応と明記されていない限り、ESDを防ぐことはできません。電動ドライバー (オプションD)は、締め付けすぎや機械的ストレスを防ぐため、繊細なコンポーネントには一般的に使用しないでください。したがって、ESDストラップは、高価なAIハードウェアの物理的な設置中に「初期不良」を防ぐための最も重要なツールです。

質問: 98

マルチノードHPLバーン中に、GPUの使用率が不均一になることがあります。どの構成がワークロードの均等な分散を保証するのでしょうか？

- A. 共有メモリアクセラレーションのためにHPL_USE_NVSHMEM=1を有効にします。
- B. HPL_RUN_GEMM_TESTS で検証をスキップする
- C. GPUノードとNUMAノードを揃えるには、--gpu-affinityと--cpu-affinityを設定します。
- D. より大きなブロックの場合は HPL_OOC_TILE_M を 8192 に変更

正解: C ([コメントを发表する](#))

マルチノードクラスタにおけるGPU利用率の不均一性は、PCIe/NUMAの不均衡の典型的な症状です。DGX H100では、各GPUセットは物理的に特定のCPUソケットと特定の高速NICセットに近接しています。ワークロード (HPL)が厳密なアフィニティなしで起動された場合、ソケットに接続されたCPUコア上で実行されるMPIランクは、0はソケット1に接続されたGPUを制御しようとする可能性があります。これにより、データはプロセッサ間リンク (UPI)を介して転送されます。

/QPI)は、直接のPCIeパスよりもレイテンシが著しく高く、帯域幅が低い。

「ボトルネック」が発生すると、一部のGPUが他のGPUよりも長くデータ待ち状態になり、結果として利用率の不均一が生じます。検証済みの解決策は、アフィニティを強制するオーケストレーションスクリプトまたはフラグを使用することです。--gpu-affinityと--cpu-affinityを設定することで、管理者は各GPUがローカルNUMAノード上のCPUコアによって管理されるようにします。この設定によりレイテンシが最小限に抑えられ、クラスタ内のすべてのGPUが同じ速度でデータを受信するため、HPLバーンインレコードを正常に完了するために必要な、フラットで高利用率のプロファイルが得られます。

質問: 99

48時間のNCCLバーンインテストにおいて、持続的なファブリックストレスを確保しつつ、サイレントデータ破損を検出するために、どのパラメータを設定すればよいのでしょうか？

- A. broadcast_perf -b 4G -e 16G -w 160

B. all_reduce_perf -b 8G -e 32G -c 1000 -z 1 -G 1000

C. all_reduce_perf -b 8G -e 32G -z 1 -G 1000

D. reduce_scatter_perf -f 2 -g 8

正解: [\(正解を表示します\)](#)

NVIDIA Collective Communications Library (NCCL) テストは、GPU クラスタの相互接続パフォーマンスを検証するためのゴールドスタンダードです。長時間のバーンイン (48 時間) では、ピーク帯域幅を測定するだけでなく、負荷がかかった状態でファブリックにストレスを与え、断続的なハードウェア障害や「サイレント データ破損」(SDC) を検出することが目的です。all_reduce_perf テストは、すべての GPU 間で双方向のデータ フローを伴うため、最も負荷の高いテストです。オプション B の特定のパラメータは重要です。-b 8G -e 32G は、400G InfiniBand リンクを飽和させる大きなバッファにメッセージ サイズ範囲を設定します。-c 1000 は、統計的有意性を確保するために多数の反復を保証します。-z 1 (check) は、数学的結果の検証を可能にするため、最も重要なフラグです。

トランシーバーの不具合により送信中にビットが反転した場合、-z 1 フラグによって不一致が検出され、エラーが報告されます。最後に、-G 1000 は、スイッチと HCA 間で熱平衡に達するのに十分な時間テストが実行されることを保証します。

質問: 100

CUDA対応MPI用に複数のGPUを搭載したサーバーを構成しています。各MPIプロセスが正しいGPUを使用するように、適切なGPUアフィニティを確保するために重要な環境変数はどれですか？

A. CUDA可視デバイス

B. CUDAデバイスオーダー

C. LDライブラリパス

D. MPI GPUサポート

E. CUDA 起動ブロック-I

正解: [A \(コメントを发表する\)](#)

CUDA_VISIBLE_DEVICES」はGPUアフィニティに不可欠です。特定のプロセスからどのGPUが見えるかを指定できます。これがないと、すべてのプロセスが同じGPUを使用しようとして、パフォーマンスのボトルネックが発生する可能性があります。GPUの列挙順序を制御します。共有ライブラリへのパスを指定します。これは仮説です。同期CUDA呼び出しを強制します。

質問: 101

Consider the following 'lspci' output snippet after installing an NVIDIA GPU:

03:00.0 VGA互換コントローラー: NVIDIA Corporationデバイス2236 (rev al) サブシステム:

Dellデバイス1234 使用中のカーネルドライバー: nvidia カーネルモジュール:

nvidiafb、nouveau、nvidia drm、nvidia この出力は何を示しており、潜在的な問題は何ですか？

A. GPUは正しくインストールされ、NVIDIA独自のドライバーがロードされています。

B. GPUはインストールされていますが、NVIDIA独自のドライバーではなくオープンソースの nouveau」ドライバーがロードされているため、パフォーマンスが最適ではない可能性があります。

C. 解決策: nouveau」をブラックリストに追加し、NVIDIA ドライバーが正しく設定されていることを確認してください。

D. デバイス2236」と表示されているように、GPUが正しく取り付けられていません。解決策 :GPUを再度取り付けてください。

E. GPUはインストールされていますが、 nvidiafb」ドライバが競合を引き起こしています。解決策 : nvidiafb」モジュールを削除してください。

F. 出力はハードウェアの故障を示しており、GPUの交換が必要です。

正解: [\(正解を表示します\)](#)

出力結果から、 nvidia」ドライバーと nouveau」ドライバーの両方がロードされていることがわかります。 nouveau」はオープンソースのドライバーであり、GPUの動作は可能になりますが、NVIDIA独自のドライバーのようなパフォーマンス上のメリットは得られません。 nouveau」が存在すると、NVIDIAドライバーの適切なロードが妨げられる場合があります。 nouveau」をブラックリストに追加することで、NVIDIAドライバーのみが使用されるようになります。

質問: 102

ベクトル加算を実行することを目的とした、以下の簡略化されたCUDAコードスニペットを考えてみましょう。

このコードがGPUハードウェアを正しく利用し、正確な結果を生成していることを検証するための重要な手順は何ですか？

A. NVIDIA CUDA Occupancy Calculator などを使用して、GPU の占有率を最大化するために、 blocksPerGrid」と threadsPerBlock」が適切に選択されていることを確認してください。

B. デバイスからホストへの cudaMemcpy」の後、 r」配列をCPUベースのベクトル加算結果と比較して正しさを検証します。

C. カーネル起動後に cudaDeviceSynchronize()」を使用して、GPU 計算が完了してからデータをホストにコピーするようにしてください。

D. メモリアクセスエラーや競合状態を検出するために、'cuda-memcheck' を使用してコードを実行します。

E. 上記のすべて

正解: [\(正解を表示します\)](#)

すべてのオプションは、重要な検証手順を強調表示します。ブロック/スレッド構成を最適化することで、GPUの利用率を最大化できます。CPUベースの計算結果と照合することで、正確性を確保できます。 cudaDeviceSynchronize()」は、データ転送前にGPU計算が完了していることを保証します。 cuda-memcheck」は、メモリエラーを検出します。これらのい

ずれかを怠ると、軽微なエラーやパフォーマンスのボトルネックが発生する可能性があります。

質問: 103

AIインフラストラクチャ向けにRoCEv2ネットワークを構成しています。RoCEv2トラフィックで一般的に使用されるUDPポート番号の範囲はどれですか？また、なぜこれを認識しておくことが重要なのでしょうか？

- A. 0-1023、これらはよく知られた港だからです。
- B. 4791、VXLAN用に予約されています。
- C. 49152-65535、動的/プライベートポート範囲、他のサービスとの競合を避けるため。
- D. 1024-49151、登録ポート範囲、一般用途向け。
- E. 問題なく任意のUDPポート番号を使用できます。

正解: ([正解を表示します](#))

RoCEv2では、一般的に動的/プライベートポート範囲 (49152~65535)が使用されます。これにより、他のサービスで使用されている既知のポートや登録済みポートとの潜在的な競合を回避できます。このポート範囲を把握することは、RoCEv2トラフィックが自由に流れるようにファイアウォールやネットワークポリシーを設定する上で非常に重要です。

質問: 104

あなたは、NVIDIA GPUを搭載したオンプレミスのデータセンターや、複数のベンダーのクラウドプラットフォームなど、多様なIT環境に展開されたAIアプリケーション間の相互運用性を確保する責任を負っています。これらの環境では、大規模なデータ処理、リアルタイム分析、機械学習モデルのトレーニングを含む複雑なAIワークフローをサポートする必要があります。一貫したパフォーマンスと柔軟性を維持するために、どの戦略を優先すべきでしょうか？

- A. ベンダーを1社に絞り、すべての環境でストレージソリューションを1つに標準化することで、管理を簡素化し、相互運用性を向上させます。
- B. 各クラウドプラットフォームでネイティブストレージソリューションのみを使用するマルチクラウド戦略を実装し、相互運用性とデータの一貫性を確保するためにミドルウェアに依存します。
- C. すべての環境で、NFSやS3などの互換性のあるストレージプロトコルとAPIを使用し、プラットフォーム間でのデータ交換と統合を容易にします。
- D. 拠点間のネットワーク帯域幅を増やすことだけに集中し、遅延を減らし、データ転送速度を向上させます。

正解: ([正解を表示します](#))

最適な戦略は、オンプレミス環境とクラウド環境全体で互換性のあるストレージプロトコルとAPIを確保することです。AIワークフローは、データ取り込み、前処理、トレーニング、チェックポイント、検証、推論、分析など、複数のステージを経ることがよくあります。各環境で互換性のないストレージインターフェイスを使用している場合、アプリケーション

ンはカスタム統合作業、データ コピー、またはワークフローの再設計が必要になる可能性があります。ファイル ベースのアクセスには NFS、オブジェクト アクセスには S3 互換 API などの共通プロトコルと API を使用することで移植性が向上し、AI ツール、データ パイプライン、オーケストレーション システムがプラットフォーム間でより一貫して動作できるようになります。ベンダーを 1 つに標準化すると管理の複雑さは軽減されますが、柔軟性が制限され、ベンダーロックインが発生する可能性があります。ミドルウェアを使用したネイティブクラウドストレージのみを使用することも可能ですが、複雑さが増し、パフォーマンスが不安定になる可能性があります。ネットワーク帯域幅を増やすとデータ転送は改善されますが、プロトコルの互換性やアプリケーションの統合は解決されません。NVIDIA AI インフラストラクチャでは、ストレージ設計は、DGX、HGX、Kubernetes、Slurm、およびクラウド接続 AI 環境全体で運用上の柔軟性を維持しながら、高スループットのデータ アクセスをサポートする必要があります。

質問: 105

新しいGPUクラスタでNCCLバーンインテストを実行する主な目的は何ですか？

- A. GPUがオペレーティングシステムによって正しく認識され、正しいドライバがインストールされているかどうかをテストします。
- B. 機械学習ワークロードにおけるGPU利用率を最大化し、ディープラーニングフレームワークを自動的に調整します。
- C. GPU通信リンクに負荷をかけることで、生産前にハードウェアまたは相互接続の問題を検出して解決します。
- D. 実際のユーザーデータと本番環境のトレーニングスクリプトを使用して、AIモデルのアプリケーション固有の実行時パフォーマンスをベンチマークします。

正解: ([正解を表示します](#))

NCCLバーンインテストの主な目的は、クラスタを本番環境にリリースする前に、GPU通信リンクに負荷をかけ、ハードウェアまたは相互接続の問題を明らかにすることです。NCCLテストでは、all-reduce、broadcast、reduce-scatter、all-gatherなどの集合的な通信パターンを実行します。これらの操作は、複数のサーバーにまたがるGPUが勾配とモデルデータを効率的に交換する必要がある分散AIトレーニングの中心となります。バーンインテストは、一定期間にわたって通信を繰り返し実行することで、不安定なケーブル、弱いリンク、スイッチの問題、HCAの問題、GPUDirect RDMAの障害、ドライバの不一致、トポロジの問題、または短時間の検証では明らかにならない断続的なエラーを明らかにするのに役立ちます。GPUの検出とドライバの可視性は、通常、NCCLバーンインではなく、nvidia-smiおよび関連するヘルスコマンドで確認されます。NCCLはディープラーニングフレームワークを自動的に調整するものではなく、実際のユーザーによるトレーニングスクリプトを使用したアプリケーションレベルのベンチマークに取って代わることを意図したものではありません。NVIDIAのAIインフラストラクチャにおいて、NCCLバーンインは、持続的なGPU通信負荷下でファブリックを検証する、製品化前の信頼性テストであり、大規模なトレーニングジョブの失敗リスクを低減します。

質問: 106

冷却能力が限られているデータセンターにGPUサーバーを設置します。パフォーマンスを著しく損なうことなく、サーバーの発熱量を最小限に抑えるには、以下のサーバー構成オプションのうちどれが最適でしょうか？すべてのオプションは互換性があると仮定します。

- A. たとえ旧世代のGPUを使うことになっても、TDP（熱設計電力）の低いGPUを選びましょう。
- B. ファンノイズと消費電力を削減するために、パッシブ冷却のCPUを使用する。
- C. BIOS/UEFI を設定して、高負荷時に CPU と GPU の周波数を積極的に制限します。
- D. GPUとCPUに液冷システムを実装する。
- E. データセンターの周囲温度を上げて、温度差を小さくします。

正解: ([正解を表示します](#))

液冷は、GPUやCPUなどの高出力コンポーネントから熱を除去する最も効果的な方法であり、過熱することなく最大限のパフォーマンスで動作させることができます。TDPの低いGPUを選択すると発熱量は減少しますが、パフォーマンスも大幅に低下します。スロットリング周波数は有効ですが、液冷は熱制約内で最適なパフォーマンスを実現します。データセンターでは冷却コストを削減する必要がありますが、サーバーの温度を下げることは直感に反します。

有効的な**NCP-AII**問題集はJPNTTest.com提供され、**NCP-AII**試験に合格することに役に立ちます！JPNTTest.comは今最新**NCP-AII**試験問題集を提供します。JPNTTest.com NCP-AII試験問題集はもう更新されました。ここで**NCP-AII**問題集のテストエンジンを手に入れます。最新版のアクセス、<https://www.jpntest.com/shiken/NCP-AII-mondaishu>
125問、30%ディスカウント、特別な割引コード: JPNshiken」

質問: 107

Your AI training pipeline involves a pre-processing step that reads data from a large HDF5 file. You notice significant delays during this step. You suspect the HDF5 file structure might be contributing to the slow read times. What optimization technique is MOST likely to improve read performance from this HDF5 file?

- A. Converting the HDF5 file to a CSV file.
- B. Storing the HDF5 file on a network file system like NFS.
- C. Reorganizing the HDF5 file to improve data contiguity and chunking.
- D. Compressing the HDF5 file using gzip.
- E. Encrypting the HDF5 file for enhanced security.

正解: ([正解を表示します](#))

Reorganizing the HDF5 file (option C) to improve data contiguity and chunking is the most effective optimization. HDF5 performance is highly dependent on how the data is laid out within the file. Contiguous data and optimal chunk sizes allow for more efficient 1/0 operations. Converting to CSV (A) loses the hierarchical structure of HDF5. Storing on NFS (B) adds network overhead. Compression (D) can reduce storage space but increases decompression overhead. Encryption (E) adds overhead without improving read performance.

質問: 108

システムエンジニアは、すべてのGPUのvGPUスケジューリング動作を、デフォルトのタイムスライス長でスケジューリングを均等に共有するように設定する必要があります。どのコマンドを使用すればよいでしょうか？

A. `esxcli system module parameters set -m nvidia -p`

`"NVreg_RegistryDwords=RmPVMRL=0x01"`

B. `esxcli グラフィックモジュールパラメータセット -m nvidia -p`

`"NVreg_RegistryDwords=RmPVMRL=0x01"`

C. `esxcli system module parameters set -m nvidia -p "NVreg_RegistryDwords=FRL=0x01"`

D. `esxcli system module parameters set -m nvidia -p`

`"NVreg_RegistryDwords=RmPVMRL=0x00"`

正解: **A** ([コメントを发表する](#))

When deploying NVIDIA vGPU on VMware ESXi, the NVIDIA driver provides several scheduling policies to determine how GPU physical resources are shared among multiple virtual machines. The default behavior is often the "Best Effort" scheduler, but for environments requiring predictable performance across all users, the "Equal Share" scheduler is preferred. This scheduler gives each vGPU an equal "time slice" of the physical GPU's engines. The configuration is managed via module parameters passed to the nvidia kernel driver during host boot. The specific registry key for this behavior is RmPVMRL. Setting RmPVMRL=0x01 enables the Equal Share scheduler (Option A). Conversely, 0x00 would revert to the default time-sliced behavior. It is critical to use system module parameters set to ensure the setting persists across reboots and is applied globally to the NVIDIA driver stack. This ensures that no single "noisy neighbor" VM can monopolize the GPU cycles, which is a common requirement in shared AI research labs or virtual desktop infrastructures where consistency is more important than raw peak throughput of a single task.

質問: 109

システム管理者がDGX H100サーバーに障害が発生していることに気づきました。再起動後、BMCのみが使用可能な状態になっています。

この行動の理由は何でしょうか？

A. ネットワークカードにリンク/接続がありません。

- B. ブートディスクに障害が発生しました。
- C. 複数のGPUが故障しました。
- D. 電源装置が2つ以上故障しています。

正解: **B** ([コメントを发表する](#))

NVIDIA DGX システムでは、ベースボード管理コントローラ (BMC) は独立したプロセッサであり、メイン CPU とオペレーティングシステムがロードに失敗した場合でも動作します。サーバーが再起動し、管理者が BMC Web インターフェイスまたは IPMI コンソールにアクセスできるにもかかわらず、OS (Ubuntu/DGX OS) がロードされない場合は、ブートディスクの障害が最も可能性の高い原因です。DGX H100 は、OS ブートボリュームに RAID-1 構成の NVMe ドライブを使用します。ミラーリングされた両方のドライブが故障した場合、またはブートパーティションが破損した場合、システムは BIOS または UEFI プロンプトでハングし、ブート可能なデバイスが見つかりません。電源ユニットの故障 (オプション D) またはネットワークリンクの障害 (オプション A) によって問題が発生する可能性はありますが、通常はそれぞれ BMC にまったくアクセスできなくなったり、リモートネットワークトラフィックが遮断されたりします。GPU の障害 (オプション C) では OS の起動は停止しません。システムは単に GPU カウントが低下した状態で起動します。したがって、BMC の「ストレージ」ログを使用してストレージの状態を確認することが正しい診断手順です。

質問: 110

A1サーバーで、GPUアクセラレーション処理中にCPU使用率が異常に高くなっています。CPUがボトルネックとなり、GPUの性能を最大限に発揮できていないかどうかをどのように判断すればよいのでしょうか？

- A. ワークロードの実行中に「top」または「htop」を使用してCPU使用率を監視します。使用率が継続的に高い場合は、ボトルネックが発生している可能性があります。
- B. 'nvidia-smi' を使用して GPU 使用率を確認してください。GPU 使用率が低く、CPU 使用率が高い場合は、CPU ボトルネックの可能性を示しています。
- C. CPUプロファイラ (例'perf')を使用してアプリケーションをプロファイリングし、CPU 負荷の高い関数を特定します。
- D. GPUワークロードと並行してCPU負荷の高いベンチマークを実行し、パフォーマンスの低下を観察します。
- E. 上記のすべて

正解: ([正解を表示します](#))

これらの方法はすべて、CPUのボトルネックを特定するための有効な手段を提供します。CPUとGPUの使用率を監視したり、アプリケーションのプロファイリングを行ったり、並列ベンチマークを実行したりすることで、CPUがGPUのパフォーマンスを制限しているかどうかを判断するのに役立ちます。

質問: 111

NVIDIA GPU を使用する Docker コンテナが断続的に「CUDA ERROR OUT OF MEMORY」エラーで失敗するという問題のトラブルシューティングを行っています。ホストシステムと個々の GPU には十分なメモリがあります。この問題は、コンテナ環境内でのメモリ割り当て方法に関連している可能性があると考えています。この問題を調査し、軽減するためにどのような手順を実行できますか？

- A. コンテナ実行時に「shm-size」フラグを使用して、コンテナの共有メモリサイズを増やします。
- B. コンテナ内の環境変数を設定して、アプリケーションから見えるGPUの数を制限します。
- C. コンテナ内の環境変数を調整して、GPUの順序が一定になるようにします。
- D. 'nvidia-smi' を使用してコンテナ内外の GPU メモリ使用量を監視し、メモリ リークや過剰な割り当てを特定します。
- E. docker run の '-compute' パラメータを使用して計算能力を下げます。

正解: ([正解を表示します](#))

コンテナ (A) 内の共有メモリが不足すると、「CUDA ERROR OUT OF MEMORY」エラーが発生する可能性があります。共有メモリのサイズを増やすと、コンテナはプロセス間通信と GPU データ転送用により多くのメモリを割り当てることができます。メモリ不足の原因を特定するには、コンテナ内外 (D) の両方で GPU メモリ使用量を監視することが重要です。「CUDA_VISIBLE_DEVICES」と (B & C) は、主に GPU の選択と順序付けに使用され、メモリ管理には使用されません。ただし、アプリケーションの設計が不十分で、必要性に関係なくすべての可視 GPU にメモリを割り当てようとする場合、GPU の可視性を制限すると、間接的に全体のメモリ消費量が削減される可能性があります。計算能力を下げてもメモリ使用量に直接影響はありませんが、アプリケーションが処理に必要なメモリは少なくなり、問題が発生する可能性があります。

質問: 112

クラスタ管理者は、UFMを使用して200ポートにわたるトランシーバのファームウェアバージョンを検証する必要があります。どのGUIベースの方法が統合ビューを提供しますか？

- A. 「デバイス」に移動し、スイッチを選択して「ケーブル」タブを開くと、ASICファームウェアとトランシーバのバージョンが表示されます。
- B. 「トポロジー」ビューを使用して、ケーブルアイコンを視覚的に検査します。
- C. 各ポートで手動で `mlxlink -d lid- <LID> -m` を実行します。
- D. すべてのスイッチログをエクスポートし、「FW Version」をgrepします。

正解: ([正解を表示します](#))

大規模な AI ファブリックを管理するには、物理層を一元的に可視化する必要があります。NVIDIA Unified Fabric Manager (UFM) は、InfiniBand ネットワーク用の包括的なダッシュボードを提供します。ファブリック全体で機能の均一性と安定性を確保するために重要なトランシーバのファームウェアを確認するには、管理者は UFM Enterprise GUI を

使用できます。「デバイス」セクションに移動して特定のスイッチを選択すると、「ケーブル」タブに、使用されているすべてのポートのテレメトリが集約されます。このビューには、トランシーバー (LinkX) またはアクティブ光ケーブル (AOC) の製造元、部品番号、および特定のファームウェアバージョンが表示されます。この統合ビューは、200 を超えるポートに対して手動で CLI クエリ (オプション C) を実行するよりもはるかに効率的です。トランシーバー全体でファームウェアを統一することで、アダプティブルーティングや輻輳制御などの最適化が 400G または 200G ファブリック全体で一貫して動作することが保証されます。

質問: 113

You need to configure persistent network settings on your BlueField SmartNIC after deploying BlueField OS. Which file should you modify to ensure these settings are applied after each reboot, assuming a Debian-based distribution?

- A. /etc/network/interfaces
- B. /etc/resolv.conf
- C. /etc/hostname
- D. /etc/sysconfig/network-scripts/ifcfg-
- E. /etc/dhcp/dhclient.conf

正解: **A** ([コメントを发表する](#))

On Debian-based systems, the '/etc/network/interfaces' file is the standard location for configuring persistent network settings. Changes made to this file will be applied on each boot. Vetc/resolv.conf is for DNS settings, &/etc/hostname' sets the system's hostname, and '/etc/sysconfig/network-scripts/ifcfg-' is more commonly found on Red Hat-based systems. letc/dhcp/dhclient.conf is for DHCP settings.

質問: 114

A critical AI model training job consistently fails on a specific GPU server in your cluster after running for approximately 24 hours.

Monitoring data shows a sudden drop in GPU power consumption followed by a system reboot. All other GPUs on the server appear normal. The server has redundant PSUs.

What is the MOST likely cause?

- A. A software bug in the A1 model causing a kernel panic specifically triggered after 24 hours of execution.
- B. Thermal runaway on the GPU due to a failing thermal interface material (TIM) between the GPU die and the heatsink.
- C. A transient power supply issue affecting only one of the redundant PSUs, triggering a system-wide protection mechanism.
- D. ECC memory errors accumulating over time, eventually leading to a non-recoverable system fault.
- E. A driver crash, causing the GPU to reset and the system to reboot.

正解: ([正解を表示します](#))

Thermal runaway (B) is the most probable cause. The 24-hour delay suggests a gradual heat buildup. A failing TIM would cause the GPU to overheat until it triggers a thermal shutdown, resulting in the power drop and reboot. While a PSU issue (C) is possible, redundant PSUs should prevent a complete failure unless one PSU is completely dead and the second PSU is overloaded by the entire load for a short period. The other options are less likely to cause this specific failure pattern.

質問: 115

あなたは、大規模な推薦システム向けの高性能ネットワークを設計する任務を負っています。このシステムでは、トレーニングと推論の両方において、低遅延かつ高スループットが求められます。クラスタ内のノードを接続するのに最も適した相互接続技術はどれでしょうか？

- A. ギガビットイーサネット
- B. 10ギガビットイーサネット
- C. InfiniBand
- D. ファイバーチャネル
- E. 100ギガビットイーサネット

正解: ([正解を表示します](#))

InfiniBandは高性能コンピューティング向けに設計されており、イーサネットやファイバーチャネルと比較して大幅に低いレイテンシと高い帯域幅を提供するため、レコメンデーションシステムなどの要求の厳しいワークロードに最適な選択肢となります。100ギガビットイーサネットは高い帯域幅を提供しますが、InfiniBandは一般的に低いレイテンシを提供します。

質問: 116

NVIDIA NVLinkスイッチの状態とパフォーマンスを監視するために一般的に使用されるコマンドラインツールはどれですか？

- A. nvidia-smi
- B. nvswitch-cli
- C. ibstat
- D. rocminfo
- E. lspci

正解: B ([コメントを发表する](#))

'nvidia-smi'はGPUに関する情報を提供しますが、'nvswitch-cli' または、スイッチの種類に応じてベンダーが提供する同様のツール)は、NVLinkスイッチ自体を管理および監視するために特化して設計されています。'ibstat'はInfiniband用、'rocminfo'はAMD GPU用、'lspci'はPCIデバイスの一覧表示用です。

質問: 117

大規模言語モデル (LLM) のトレーニングジョブが、クラスタ内の複数の NVIDIA AI 100 GPU で実行されています。単一サーバー内の GPU は NVLink を介して効率的に通信していますが、サーバー間のイーサネット通信がボトルネックになりつつあります。ケーブルとトランシーバーの選択に焦点を当てた以下の戦略のうち、このサーバー間通信のボトルネックを最も効果的に解消できるのはどれでしょうか？ 2つ選択してください)

- A. 既存の Cat6 イーサネットケーブルを Cat8 ケーブルに交換します。
- B. 適切なトランシーバーと光ファイバーケーブルを使用して、サーバー間の接続を利用可能な最高イーサネット速度 (例: 100GbE から 400GbE) にアップグレードします。
- C. 適切な InfiniBand ケーブルとトランシーバーを使用して、サーバー間通信の相互接続技術として InfiniBand を実装します。
- D. LLM トレーニングジョブのバッチサイズを縮小します。
- E. 既存のトランシーバーをすべて アクティブ光ケーブル (AOC) に交換します。

正解: [\(正解を表示します\)](#)

適切なトランシーバーと光ファイバーを使用してイーサネットの速度を高速化 (例: 400GbE) すると、サーバー間通信の帯域幅が大幅に増加します。高性能コンピューティング相互接続用に設計された技術である InfiniBand を導入すると、イーサネットよりもさらに優れたパフォーマンスが得られます。Cat8 ケーブルはイーサネットの信号品質を向上させますが、帯域幅を根本的に増加させるわけではありません。バッチサイズを小さくすると転送されるデータ量は減りますが、トレーニングの収束に影響を与える可能性があります。AOC は信号品質を向上させることができますが、より高速なイーサネットや InfiniBand ほど効果的に根本的な帯域幅の制限に対処することはできません。

質問: 118

If two ports must be connected, but one is SFP and one is QSFP, for example, to connect a 25 GbE HOST CHANNEL ADAPTER to a QSFP port capable of both 100 GbE and 25 GbE, which of the following solutions would best meet this requirement?

- A. SFP コネクタ
- B. SFP - 1G BASE-T (RJ45) アダプタ
- C. QSA アダプター

正解: [C \(コメントを发表する\)](#)

QSA (QSFP - SFP アダプタ) は、シングルレーンの SFP/SFP28 トランシーバー (通常 10G または 25G) を 4 レーンの QSFP/QSFP28 スイッチポートに接続できるようにする機械的および電気的なブリッジです。AI インフラストラクチャでは、100G/200G ポート全体を無駄にしたり、専用のブレークアウトケーブルを必要とせずに、低速の管理サーバーやレガシーノードを高速バックボーンスイッチに接続するためによく使用されます。QSA アダプタは、SFP モジュールのシングルレーンを QSFP ポートの最初のレーンにマッピングします。これは、リンクの信号完全性と遅延特性を維持する「パススルー」ソリューションです。NVIDIA ネットワーク環境におけるポート密度の不一致に対する検証済みのハードウェアソリューションです。

質問: 119

数日間にわたるNeMoのバーンインテスト中に、断続的に「GPUがバスから切断されました」というエラーが発生します。ハードウェア障害を特定するための診断方法はどれですか？

- A. 代替メモリ共有のために HPL_USE_NVSHMEM を有効にします。
- B. バーンインと同時にDCGM診断を実行して、GPUの状態指標を監視します。
- C. 計算を簡素化するために、BERT モデルから GPT モデルに切り替えます。
- D. メモリ負荷を軽減するためにブロックサイズを500MBに縮小します。

正解: [\(正解を表示します\)](#)

エラー「GPU fell off bus」は、GPUとCPU/PCIeスイッチ間のPCIeリンクが切断された重大な障害であり、多くの場合、熱ストレス、電源の不安定性、またはハードウェアの物理的な欠陥が原因です。NVIDIA NeMo（大規模言語モデルフレームワーク）のような負荷の高いワークロード中に根本原因を特定するには、管理者は高精度のテレメトリを収集する必要があります。DCGM（データセンターGPUマネージャ）の診断機能は、まさにこのシナリオのために設計されています。ワークロードと同時に`dcgmi diag -r 3`（包括的なハードウェアストレステスト）を実行するか、`dcgmi health --check`でヘルス状態を監視することで、システムはPCIeリプレイカウント、温度スパイク、XIDエラーなどのパラメータが発生した正確な瞬間をキャプチャできます。このデータにより、エンジニアは特定のH100モジュールに問題があるのか、それとも問題がシステム全体（たとえば、マザーボード上のPCIeスイッチの故障）にあるのかを判断できます。ワークロードを下げる（オプションCまたはD）ことで症状を隠せるかもしれませんが、ハードウェアがピーク電力とデータスループットを処理できないという根本的な原因を診断することはできません。

質問: 120

複数のNVIDIA A100 GPUを搭載したサーバーラックの冷却戦略を計画する際に、最も重要な考慮事項は次のうちどれですか？

- A. サーバー室の温度を摂氏25度以下に保ってください。
- B. ラックから熱気を効率的に排出し、冷気を吸い込むように空気の流れを最適化します。
- C. CPUには液冷を使用し、GPUには空冷を使用する。
- D. GPUメモリチップにサーマルペーストを塗布します。
- E. Increasing the fan speed of the server chassis fans to maximum.

正解: [B \(コメントを発表する\)](#)

While ambient temperature is important, optimized airflow is crucial for removing the heat generated by the GPUs. Focusing solely on CPU cooling neglects the GPU heat. Applying thermal paste to memory chips is generally unnecessary unless specifically recommended by the manufacturer. Maximizing fan speed can help, but efficient airflow design is more effective.

質問: 121

マルチテナントKubernetesクラスタでGPUワークロードを分離する必要があるシナリオを考えてみましょう。リソースレベルとネットワークレベルの両方で強力な分離を実現するのに最も適したKubernetes構成は次のうちどれでしょうか？

- A. リソースクォータとネットワークポリシーを使用した名前空間の使用。
- B. ラベルとセレクタを使用して、特定のGPUノードにワークロードをスケジュールします。
- C. ティントとトレランスを使用して、GPUノードを特定のワークロードに割り当てます。
- D. ポッドの親和性と反親和性のルールを使用してポッドの配置を制御します。
- E. ノードアフィニティのみを使用します。

正解: ([正解を表示します](#))

名前空間は、Kubernetes クラスタ内で論理的な分離を提供します。リソース クォータは、名前空間が消費できるリソース (GPU を含む) を制限し、ネットワーク ポリシーは名前空間間のネットワーク トラフィックを制御することで、強力な分離を実現します。オプション B、C、D、および E は、Pod の配置をある程度制御できますが、リソース クォータとネットワーク ポリシーを備えた名前空間ほど、リソースとネットワークの分離レベルは高くありません。

有効的な**NCP-AII**問題集はJPNTTest.com提供され、**NCP-AII**試験に合格することに役に立ちます！JPNTTest.comは今最新**NCP-AII**試験問題集を提供します。JPNTTest.com NCP-AII試験問題集はもう更新されました。ここで**NCP-AII**問題集のテストエンジンを手に入れます。最新版のアクセス、<https://www.jpntest.com/shiken/NCP-All-mondaishu> **125問、30%ディスカウント**、特別な割引コード: **JPNshiken**」

質問: 122

NVIDIA GPUとNVLinkを使用するデータセンターにおいて、ネットワーク経由でトラフィックをルーティングする場合と比較して、NVLinkを直接接続するネットワークトポロジーを使用する主な利点は何ですか？

- A. ネットワークセキュリティの強化
- B. GPU間の帯域幅の向上とレイテンシの低減
- C. ネットワークインフラのコスト削減
- D. 簡略化されたネットワーク構成
- E. 電力効率の向上

正解: **B** ([コメントを发表する](#))

NVLinkを直接接続することで、従来のネットワーク経由でトラフィックをルーティングする場合と比べて、帯域幅が大幅に向上し、レイテンシも低減されます。これは、大規模なAIトレーニングなど、GPU間の通信を頻繁に行う必要があるアプリケーションにとって非常

に重要です。NVLinkを直接接続することで設定が簡素化される場合もありますが、最大の利点はパフォーマンスの向上です。

質問: 123

NVIDIA Spectrum-XスイッチをA1インフラストラクチャで使用する場合、従来のイーサネットスイッチと比較して、以下のどの点が主な利点となりますか？ (3つ選択してください)

- A. 港湾あたりのコストが低い。
- B. RoCE (RDMA over Converged Ethernet) およびInfiniBandプロトコルをサポートし、高帯域幅、低遅延の通信を実現します。
- C. ネットワークパフォーマンス最適化のための高度なテレメトリおよび監視機能。
- D. 分散A1トレーニングで使用される集団通信操作のためのハードウェアベースのアクセラレーション。
- E. IPv6のネイティブサポート。

正解: **B,C,D** ([コメントを发表する](#))

Spectrum-Xスイッチは、高性能コンピューティングおよびA1ワークロード向けに設計されています。低遅延通信のためのRoCEおよびInfiniBandをサポートし、ネットワーク最適化のための高度なテレメトリ機能を提供するとともに、集合通信操作のためのハードウェアベースのアクセラレーションを搭載しており、分散型A1トレーニングの効率を向上させます。Spectrum-XはIPv6をサポートしていますが、これは最新のイーサネットスイッチでも一般的な機能です。Spectrum-Xスイッチは、その高度な機能と性能のため、一般的に基本的なイーサネットスイッチに比べてポートあたりのコストが高くなります。

質問: 124

デュアル電源を搭載したDGX A100サーバーで、BMCログに重大な電源イベントが報告されました。一方の電源ユニットは「劣化」状態を示していますが、もう一方は正常に見えます。継続的な運用を確保し、データ損失を防ぐために、どのような緊急措置を講じるべきでしょうか？

- A. 故障したPSUへのさらなる損傷を防ぐため、サーバーを直ちに正常にシャットダウンしてください。
- B. 劣化した電源ユニットを交換用ユニットとホットスワップします。
- C. 残りの電源ユニットの負荷と温度を注意深く監視し、安定している場合は、予定されているメンテナンス期間まで運転を継続します。
- D. 'nvidia-smi' を使用して GPU の電力制限を下げ、サーバー全体の消費電力を削減します。
- E. 潜在的な電源ユニットの完全な故障の影響を最小限に抑えるため、すべてのワークロードをクラスタ内の他のサーバーに移行します。

正解: **B,E** ([コメントを发表する](#))

劣化した電源ユニット B)をホットスワップすることで冗長性が回復します。ワークロードを移行することで E)、残りの電源ユニットが故障した場合のデータ損失やサービス中断のリスクを最小限に抑えることができます。ホットスワップが可能な場合、サーバーをシャットダウンすると A)、不要なダウンタイムが発生します。残りの電源ユニット C)を監視することは良い習慣ですが、冗長性の回復やリスク軽減の代替にはなりません。GPUの電力制限を下げる D)ことで、さらなる負荷を防ぐことができますが、これは一時的な解決策であり、パフォーマンスに影響を与えます。

質問: 125

When setting up a multi-server, multi-GPU environment using NVLink switches, what is the primary consideration when planning the network topology for optimal performance?

- A. Minimizing the number of hops between GPUs that need to communicate frequently.
- B. Maximizing the distance between servers to improve cooling.
- C. Using a star topology for simplified management.
- D. Ensuring all servers are on the same subnet for ease of configuration.
- E. Placing servers near the network's edge to reduce latency.

正解: [A \(コメントを发表する\)](#)

The primary goal is to reduce latency and maximize bandwidth. Minimizing hops between frequently communicating GPUs is essential for achieving this. Server placement for cooling is a secondary concern, and topology choice should prioritize performance over simplicity. Subnet configuration and server placement relative to the network edge are less critical for NVLink-specific performance.

質問: 126

AIインフラストラクチャでは、ノード間通信にNVIDIA ConnectX-7 NICを使用しています。要件は、100メートルの距離で低遅延かつ400GbEの帯域幅を実現することです。このシナリオに最も適したトランシーバーとケーブルの組み合わせはどれでしょうか？

- A. OM3マルチモード光ファイバーケーブル付きQSFP-DD SR8トランシーバー。
- B. シングルモード光ファイバーケーブル付きQSFP-DD DR4トランシーバー。
- C. OM4マルチモード光ファイバーケーブル付きQSFP-DD LR8トランシーバー。
- D. OM4マルチモード光ファイバーケーブル付きQSFP-DD SR4トランシーバー。
- E. シングルモード光ファイバーケーブル付きQSFP-DD FR4トランシーバー。

正解: [\(正解を表示します\)](#)

QSFP-DD SR4とOM4を組み合わせると、マルチモードファイバーを使用して短距離（最大100m）で400GbEを実現します。DR4とFR4はシングルモードファイバーを必要とし、通常はより長距離で使用されます。LR8は通常、指定された距離でシングルモードファイバーを必要とします。SR8はOM4/OM5との組み合わせで最適に動作するため、OM3とSR8を組み合わせても400GbEを実現することは難しいでしょう。

質問: 127

BlueField-3 DPUは、制御プレーンとデータプレーンの両方の機能を実行するように構成されています。最近のソフトウェアアップデート後、データプレーンのパフォーマンスが著しく低下しているにもかかわらず、制御プレーンは正常に動作していることに気づきました。アップデートによってコードのバグが発生していないと仮定した場合、最も可能性の高い原因は何ですか？また、この問題を診断するための最善のアプローチは何ですか？

- A. リソース競合; 'perf' または 'bpfftrace' を使用してデータプレーンプロセスをプロファイリングし、リソースのボトルネック (CPU、メモリ、キャッシュ) を特定します。
- B. ドライバーの互換性の問題。Mellanox OFED ドライバーを以前のバージョンにダウングレードしてください。
- C. ファームウェアの破損。BlueField DPIJ に最新のファームウェアイメージを再フラッシュしてください。
- D. ネットワーク設定ミス。ネットワークインターフェースのMTUとQoS設定を確認してください。
- E. 電力制限; BMC を介して DPU の消費電力と熱状態を確認します。

正解: ([正解を表示します](#))

コードのバグがないと仮定すると、リソース競合が最も可能性の高い原因です。アップデートによって制御プレーンまたはデータプレーンのリソース要求が増加し、競合が発生した可能性があります。「perf」または「bpfftrace」を使用してデータプレーンプロセスをプロファイリングすると、ボトルネックを特定するのに役立ちます。リソース競合が原因ではないことを確認した後、ドライバのダウングレードやファームウェアの再フラッシュは、より抜本的な対策となります。

質問: 128

Kubernetes環境で、ノードごとに複数のGPUを搭載した状態でNVIDIA Container Toolkitを使用しています。Podが「任意の」GPUを要求するのではなく、ノード上の特定のGPUを要求できるようにしたいと考えています。NVIDIA Device Pluginと連携して、このようなきめ細かなGPUリソース割り当てを実現するには、どのKubernetes機能を使用すればよいでしょうか？

- A. 資源割当
- B. ノードアフィニティ
- C. トポロジーマネージャ
- D. デバイスプラグインAPI
- E. 汚染と寛容

正解: ([正解を表示します](#))

Kubernetes トポロジーマネージャ (C) を使用すると、リソース割り当て (GPU を含む) を特定の NIMMA ノードに合わせることができます。これは、ノードごとに複数の GPU を扱う場合のパフォーマンスにとって非常に重要です。リソース クォータ (A) はリソースの使用を制限しますが、特定の GPU の選択は制御しません。ノード アフィニティ (B) は、特定の GPU ではなく、ラベルに基づいてノードを選択します。デバイス プラグイン API (D) は

GPIIの検出を可能にしますが、ノード内でのきめ細かな割り当てにはトポロジマネージャが必要です。テイントとトレランス (E) は、対応するトレランスがない限り、特定のノードに Pod がスケジューラれないようにするために使用され、特定の GPU を直接選択することはできません。

質問: 129

サーバーに新しくインストールされた NVIDIA AIOO Tensor Core GPU の検証を担当していただきます。GPU が正しく認識され、期待されるパフォーマンスレベルで動作していることを確認する必要があります。以下の項目を検証するための手順 (コマンドとツールを含む) を説明してください。1) GPU の存在と基本情報、2) PCIe 帯域幅とリンク速度、3) 負荷がかかった状態での持続的な計算パフォーマンス。

A. 1) NVIDIA の存在を確認するには「`lspci | grep NVIDIA`」を、基本情報を確認するには「`nvidia-smi`」を使用します。2) 帯域幅/速度を確認するには「`nvidia-smi -q -d pcie`」を使用します。3) TensorFlow ResNet50 ベンチマークを実行します。

B. 1) 'nvidia-smi' を使用して、存在情報と基本情報を取得します。2) PCIe の速度は関係ありません。3) CUDA アプリケーションの実行中に 'nvprof プロファイラ' を実行します。

C. 1) BIOS 設定で GPU 検出を確認します。2) 'lspci -vv' を使用して PCIe の速度を確認します。3) PyTorch ImageNet トレーニングスクリプトを実行します。

D. 1) 'nvidia-smi' を使用して、存在情報と基本情報を確認します。2) 'nvidia-smi -q -d pcie' を使用して、帯域幅/速度を確認します。3) CUDA ベースの行列乗算ベンチマーク (cuBLAS など) を、行列サイズを増やしながらか実行し、パフォーマンスを監視します。

E. 1) 'nvidia-smi' を使用して、存在情報と基本情報を取得します。2) 'nvlm-link-monitor' を使用して、帯域幅/速度を取得します。3) GPU のボトルネックを回避するために、CPU に負荷のかかるベンチマークを実行します。

正解: ([正解を表示します](#))

NVIDIA GPU の情報を取得するには、「`nvidia-smi`」が主要なツールです。「`nvidia-smi -q -d pcie`」を実行すると、PCIe の詳細が表示されます。CUDA ベースのベンチマークは、GPU のパフォーマンスを分離して評価します。他のオプションにも一定の精度はありますが、完全性や最適なターゲット設定には欠けます (例 ResNet50 は他のフレームワークに依存しています)。CPU バウンドのベンチマークでは、GPU の性能を十分にテストすることはできません。

質問: 130

Consider a scenario where you are using NCCL (NVIDIA Collective Communications Library) for multi-GPU training across multiple servers connected via NVLink switches. Which NCCL environment variable would you use to specify the network interface to be used for communication?

A. NCCL PORT

B. NCCL SOCKET IFNAME

C. NCCL NET INTERFACE

D. NCCL 1B HCA

E. NCCL COMM ID

正解: ([正解を表示します](#))

is the correct environment variable to specify the network interface used by NCCL. Is for Infiniband, and the other options are not directly related to specifying the network interface.

質問: 131

DGX A100クラスタをHPL-AI 2.0にアップグレードしたところ、パフォーマンスが2倍向上しました。この改善の主な要因はどの最適化でしょうか？

A. 計算を高速化するために問題サイズ (N) を縮小する。

B. MPI対応GPU通信により、CPUのボトルネックとGPUのアイドル時間を削減します。

C. ファームウェアのアップデートと関連する設定により、GPUクロック速度を2倍にします。

D. ドライバのアップデートによるNVLink帯域幅の自動倍増。

正解: B ([コメントを发表する](#))

HPL-AI (High-Performance Linpack for AI) は、反復的な改良によって FP64 の精度を維持しながら、低精度の演算 (FP16/BF16/TF32) を利用する点で、従来の HPL とは異なります。バージョン 2.0 以降で見られるパフォーマンスの大幅な向上は、主に通信レイヤーの進歩によるものです。マルチノード DGX クラスタでは、MPI (Message Passing Interface) ランクの管理や InfiniBand ファブリックを介した GPU 間のデータ転送の調整を行う際に、CPU がボトルネックになることがよくあります。

MPI対応のGPU通信を実装し (SharpV3やNCCL統合などの技術を活用)、ベンチマークは「CPUが次の計算ブロックをオーケストレーションするのを待つ GPUアイドル時間」を大幅に削減します。この最適化により、集中的な行列乗算フェーズ中もGPUの利用率がほぼ100%に維持されます。集合演算をオフロードし、GPUDirect RDMAを使用することで、システムはホストメモリとCPUをバイパスし、従来のCPU負荷の高い協調方式と比較して、ベンチマークの実効スループットを実質的に2倍に向上させます。

有効的な**NCP-AII**問題集はJPNTTest.com提供され、**NCP-AII**試験に合格することに役に立ちます！JPNTTest.comは今最新**NCP-AII**試験問題集を提供します。JPNTTest.com NCP-AII試験問題集はもう更新されました。ここで**NCP-AII**問題集のテストエンジンを手に入れます。最新版のアクセス、<https://www.jpntest.com/shiken/NCP-All-mondaishu>
125問、30%ディスカウント、特別な割引コード: **JPNshiken**」