



GPU カードの取り付け

この章は次のトピックで構成されています。

- [サーバファームウェアの要件 \(1 ページ\)](#)
- [GPU カードの構成規則 \(2 ページ\)](#)
- [すべての GPU に関する要件：メモリマップド I/O 4 GB 以上 \(5 ページ\)](#)
- [倍幅 GPU カードの取り付け \(6 ページ\)](#)
- [ヒートシンクの交換 \(9 ページ\)](#)
- [GPU エアブロッカーの交換 \(14 ページ\)](#)
- [NVIDIA ライセンスポータルからライセンスサーバへのグリッドライセンスのインストール \(20 ページ\)](#)

サーバファームウェアの要件

次の表に、サポートされている GPU カードのサーバファームウェアの最小バージョンを示します。

GPU カード	Cisco IMC/BIOS の最低限必要なバージョン
Intel Flex 140 PCIe、75W、Gen4 x8	4.1(3)
Intel Flex 170 PCIe、150W、Gen4 x16	4.1(3)
NVIDIA Tesla A100、80GB、300W、パッシブ (UCSC-GPUA100-80-D または HX-GPU-A100-80-D)	4.1(3)
NVIDIA Tesla A40 RTX、48GB、300 W、パッシブ (UCSC-GPU-A40-D)	4.1(3)
NVIDIA Tesla A30、24GB、180 W、パッシブ (UCSC-GPU-A30-D)	4.1(3)
NVIDIA A16 PCIe、250W、64GB (4x16GB) 、 (UCSC-GPU-A16-D)	4.1(3)

GPU カード	Cisco IMC/BIOS の最低限必要なバージョン
NVIDIA H100 PCIe、350W、Gen 5 x16、(UCSC-GPU-H100-D)	4.1(3)
NVIDIA L4 PCIe、72W、Gen 4 x16 (UCSC-GPU-L4-D)	4.1(3)
NVIDIA L40 PCIe、300W、Gen 4 x16、(UCSC-GPU-L40-D)	4.1(3)

GPU カードの構成規則

GPU カードを使用して、サーバを設定するときは、次の規則に注意してください。

- UCSC-C240-M7SX および UCSC-C240-M7SN サーバは、後で GPU を受け入れるようにサーバをプリセットする「GPU 対応」設定をサポートします。

この構成では、ロープロファイルヒートシンク (UCSC-HSLP-C220M7) と、一部の個々のライザーの中央スロット (ライザー 2 スロット 4、およびライザー 3 スロット 8) (UCSC-RISAB-24XM7) に取り付けられる GPU エアブロッカーを備えたサーバを設定します。GPU エアブロッカーは、それが取り付けられているスロットとその上のスロットをブロックする 2 倍幅の部品で、気流と換気を適切に制御します。

GPU エアブロッカーは GPU 対応構成に事前に取り付けられているため、サーバのコンピューティング能力を拡張することを選択したときに GPU のみが必要になります。GPU 対応設定には、次の考慮事項があります。

- サーバは、現在 GPU が取り付けられていない場合でも、GPU がインストールされているサーバと同じ温度制限に従う必要があります。GPU が存在しない場合でも、温度制限に従っておけば、後で GPU を取り付けるときに正しく動作します。
- この構成には、薄型ヒートシンクと GPU エアブロッカーが必要です。GPU 対応設定を注文する場合は、シスコのオンライン注文および設定ツールを使用してサーバを注文するときに、GPU エアブロッカー PID を選択して GPU 対応設定を有効にする必要があります。ツールに表示される追加のルールに従います。
- エアブロッカーには 2 つのバージョンがあります。1 つは、Sapphire Rapids CPU を搭載したシステムと、定格消費電力が 75W 未満の GPU を搭載したサーバ用です。1 つは、Emerald Rapids CPU と、定格電力消費が 75W を超える 1 つ以上の GPU を搭載したサーバ用です。これらの GPU エアブロッカーの詳細については、「[GPU エアブロッカーの交換 \(14 ページ\)](#)」を参照してください。
- GPU エアブロッカーは、GPU 構成サーバまたは GPU 対応サーバの空の GPU スロットに必要です。
 - これらのサーバでは、必要に応じて工場での GPU エアブロッカーが取り付けられています。

- ただし、GPU スロットから NIC または GPU を取り外す場合は、適切なエアフローを確保するためにエア ブロッカーを取り付ける必要があります。
- CIMC および UCSM 管理では固有の SBIOS ID が必要になるため、GPU カードはすべてシスコから購入してください。
- サーバで異なるブランドまたはモデルの GPU カードを混在させないでください。
- GPU は、ライザー 1B またはライザー 3B ではサポートされていません。ライザー 3B は、GPU を機械的に受け入れることができません。
- UCSC-C240M7SX および UCSC-C240M7SN サーバーは、最大 300W、PCIe Gen4 速度のライザー 3C で1つのフルハイト、フルレングス、ダブル幅 GPU (PCIe スロット7のみ) をサポートします。
ライザー 1A と 2A はどちらも、最大 300W および PCIe Gen4 速度のフルハイト、フルレングス、ダブル幅の GPU をサポートできます。
ライザー 1C および 2C は、最大 350W で PCIe Gen5 速度のフルハイト、フルレングス、ダブル幅の GPU をサポートできます。
- 次の表に、サポートされているさまざまな GPU の詳細を示します。

GPU	GPU 情報	ライザーおよび設置に関する注意事項
Intel Flex 140	HHHL、75W、PCIe Gen 4 x8	Gen 4 または Gen 5 ライザーの両方、最大 5 つの GPU をサポート <ul style="list-style-type: none"> • Gen 4 ライザーでは、GPU はスロット 2、3、5、6、および 7 (ライザー 3C) でサポートされます。 • Gen 5 ライザーでは、GPU はスロット 1、2、4、および 5 (ライザー 1C および 2C) でサポートされます。また、スロット 7 (Gen 4 ライザー 3C)

GPU	GPU 情報	ライザーおよび設置に関する注意事項
Intel Flex 170	FHFL、シングル幅 GPU、150W、PCIe Gen 4 x16	<p>Gen 4 ライザー（最大 5 つの GPU をサポート）または Gen 5 ライザー（最大 3 つの GPU をサポート）の両方。</p> <ul style="list-style-type: none"> • Gen 4 ライザーでは、GPU はスロット 2、5、および 7（ライザー 3C）でサポートされます。 • Gen 5 ライザーでは、GPU はスロット 2 および 5（ライザー 1C および 2C）でサポートされます。また、スロット 7（Gen 4 ライザー 3C） <p>電源ケーブルが必要 (UCS-M10CBL-C240M5)</p>
Nvidia H100	FHFL、ダブル幅 GPU、350W、PCIe Gen 5 x16	<p>Gen 5 ライザーのみ、スロット 2 および 5 で最大 2 つの GPU をサポートします。</p> <p>電源ケーブルが必要 (UCS-G5GPU-C240M7)</p>
Nvidia L4	HHHL、72W、PCIe Gen 4 x16	<p>Gen 4 ライザー（最大 8 個の GPU をサポート）または Gen 5 ライザー（最大 5 個の GPU をサポート）の両方。</p> <ul style="list-style-type: none"> • Gen 4 ライザーでは、GPU はすべてのスロットでサポートされます（ライザー 1A、2A、および 3A）。 • Gen 5 ライザーでは、GPU はスロット 1、2、4、および 5（ライザー 1C および 2C）でサポートされます。また、スロット 7（Gen 4 ライザー 3C）

GPU	GPU 情報	ライザーおよび設置に関する注意事項
Nvidia L40	FHFL、ダブル幅 GPU、300W、PCIe Gen 4 x16	<p>Gen 4 ライザー（最大3つの GPU をサポート）または Gen 5 ライザー（最大3つの GPU が、Gen 5 ライザーの 2 GPU と Gen 4 ライザー 3 の 1 つの GPU としてサポート）の両方。</p> <ul style="list-style-type: none"> • Gen 4 ライザーでは、GPU はスロット 2、5、および 7 でサポートされます。 • Gen 5 ライザーでは、GPU はスロット 2、5、および 7 でサポートされます。 <p>電源ケーブルが必要 (CBL-L40GPU-C240M7)</p>

- UCS 電力計算ツール (<http://ucspowercalc.cisco.com>) を使用して、サーバ構成に基づいて必要な電源を確認してください。

すべての GPU に関する要件：メモリマップド I/O 4 GB 以上

サポートされているすべての GPU カードで、4 GB 以上のメモリ マップド I/O (MMIO) を許可する BIOS 設定の有効化が必要です。

- スタンドアロン サーバ：サーバをスタンドアロン モードで使用する場合、この BIOS 設定はデフォルトで有効です。

[Advanced] > [PCI Configuration] > [Memory Mapped I/O Above 4 GB] を **[Enabled]** にする

ブート中にメッセージが表示されたら、**F2** キーを押して BIOS Setup ユーティリティに切り替えます。

- サーバが Cisco UCS Manager と統合されてサービス プロファイルによって制御され、かつ GPU が存在する場合、この設定はサービス プロファイルでデフォルトで有効になっています。

この設定を手動で変更するには、次の手順を実行します。

ステップ 1 サービス プロファイルの設定方法については、以下からご使用のリリースの GUI または CLI の Cisco UCS Manager コンフィギュレーションガイドを参照してください。

[Cisco UCS Manager コンフィギュレーションガイド](#)

ステップ 2 サーバ関連ポリシーの設定 > BIOS 設定の構成の章を参照してください。

ステップ 3 プロファイルの PCI 構成 BIOS 設定のセクションで、[Memory Mapped IO Above 4GB Config] を以下のいずれかに設定します。

- **[Disabled]** : 64 ビット PCI デバイスを 64 GB 以上のアドレス空間にマッピングしません。
- **Enabled** : 64 ビット PCI デバイスの I/O を 64 GB 以上のアドレス空間にマッピングします。
- **[Platform Default]** : ポリシーで、サーバの BIOS デフォルト値に含まれるこの属性の値が使用されます。これは、この項目にデフォルトの [Enabled] 設定を使用するようにサーバ BIOS が設定されているとわかっている場合のみ使用します。

ステップ 4 サーバをリブートします。

- (注) Cisco UCS Manager は、BIOS ポリシーまたはデフォルトの BIOS 設定を通じて、Cisco Integrated Management Controller (CIMC) バッファに BIOS 設定の変更をプッシュします。これらの変更はバッファ内にとどまり、サーバがリブートされるまでは有効になりません。

倍幅 GPU カードの取り付け

次の手順に従って、NVIDIA 倍幅 GPU を取り付けまたは交換します。

Cisco IMC バージョン 4.3(1) 以降では、サーバは最大 3 つの NVIDIA GPU をサポートできません。サポートされている GPU のリストについては、[サーバファームウェアの要件 \(1 ページ\)](#) を参照してください。

次の表は、PCIe HDD および SSD を搭載したサーバの周囲温度のしきい値を示しています。

表 1: PCIe サーバ周囲温度

SKUの詳細	ストレージハードウェア オプション	
	すべての NVMe またはすべての SAS ストレージ	すべての NVMe またはすべての SAS Plus 4 リア HDD
XCC、350 W、Gen 4 または Gen 5	35 C 通常の周囲温度 T	30 C の通常の周囲温度 T
MCC、300 W Gen 4 または Gen 5		

次の表は、PCIe SSD および GPU を備えたサーバの周囲温度のしきい値を示しています。

表 2: PCIe サーバ周囲温度

SKUの詳細	ストレージハードウェアオプション	
	すべての NVMe ストレージと GPU	4つの NVMe SSD と 20 の SAS HDD と GPU
XCC、350 W、Gen 4 または Gen 5	30 C の通常の周囲温度 T	30 C の通常の周囲温度 T
MCC、300 W Gen 4 または Gen 5		

NVIDIA GPU カードには、2本の電源ケーブル（ストレートケーブルと Y 字型ケーブル）が同梱されている場合があります。ストレートケーブルは、このサーバで GPU カードに電源を接続するために使用します。Y 字型ケーブルは使用しないでください。Y 字型は、外部デバイスに GPU カードを接続するためにのみ使用します。

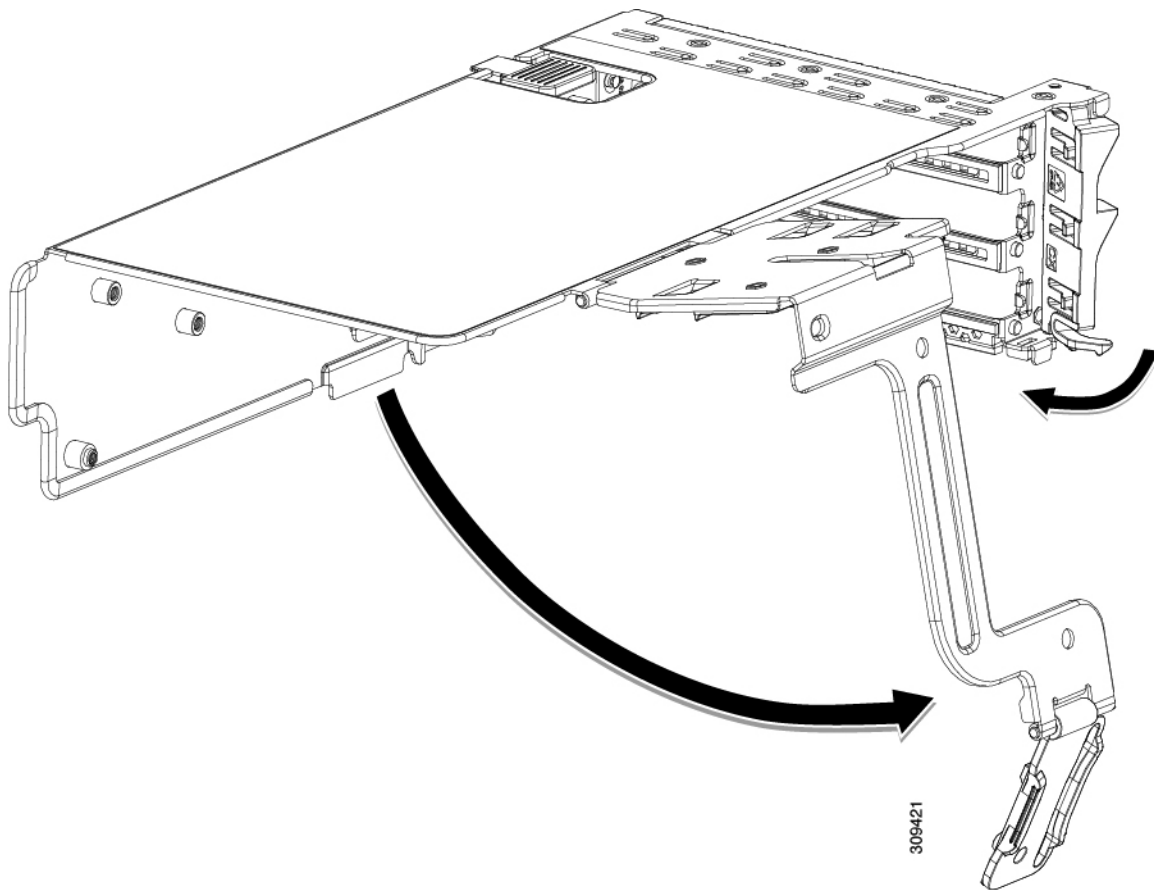


注意 GPU エア ブロッカーは、GPU 構成または GPU 対応サーバの空の GPU スロットに取り付ける必要があります。空の GPU スロットがあるままでサーバを動作させないでください。

サポートされる NVIDIA GPU には、C240 M5 NVIDIA ケーブル（UCS-P100CBL-240M5）が必要です。

- ステップ 1** サーバのシャットダウンと電源切断の説明に従って、サーバをシャットダウンして電源を切ります。
- ステップ 2** 上部カバーを取り外せるように、サーバをラックの前方に引き出します。場合によっては、リアパネルからケーブルを取り外して隙間を空ける必要があります。
- 注意** コンポーネントへの接触や確認作業を安全に行えない場合は、ラックからサーバを取り出してください。
- ステップ 3** サーバ上部カバーの取り外しの説明に従ってサーバから上部カバーを取り外します。
- ステップ 4** 既存の GPU カードを取り外します。
- GPU カードから既存のケーブルを外します。
 - 両手で PCIe ライザーの金属製ブラケットを持ってまっすぐ持ち上げ、マザーボード上のソケットからコネクタを外します。ライザーを静電気防止素材の上に置きます。
 - ライザーの底部で、固定プレートを留めているクリップを押し下げます。
 - ヒンジ付き固定プレートを回して開け、中にアクセスできるようにします。
 - カードの背面パネルタブを固定しているヒンジ付きプラスチック製固定具を開きます。
 - PCIe ライザーの電源コネクタから GPU カードの電源ケーブルを外します。
 - GPU カードの両端を均等に引いて、PCIe ライザーのソケットから取り外します。

図 1: PCIe ライザー カードの固定機構



1	ヒンジ付き固定プレートのリリース ラッチ	3	ヒンジ付きカードタブ固定具
2	ヒンジ付き固定プレート	-	

ステップ 5 新しい GPU カードを取り付けます。

(注) [GPU カードの構成規則 \(2 ページ\)](#) で説明されているこのサーバの設定ルールに従います。

- a) GPU カードをライザーのソケットの位置に合わせ、カードのエッジコネクタをソケットにゆっくと押し込みます。コネクタに支障をきたすことを防ぐためにカードの両隅を均等に押します。
- b) GPU の電源コードを接続します。ストレート電源ケーブルコネクタは色分けされています。ケーブルの黒いコネクタを GPU カードの黒いコネクタに、ケーブルの白いコネクタを PCIe ライザーの白い GPU POWER コネクタに接続します。

注意 ストレート電源ケーブルを逆向きに接続しないでください。ケーブルの黒いコネクタは GPU カードの黒いコネクタに接続します。コードの白いコネクタを PCIe ライザーの白いコネクタに接続します。

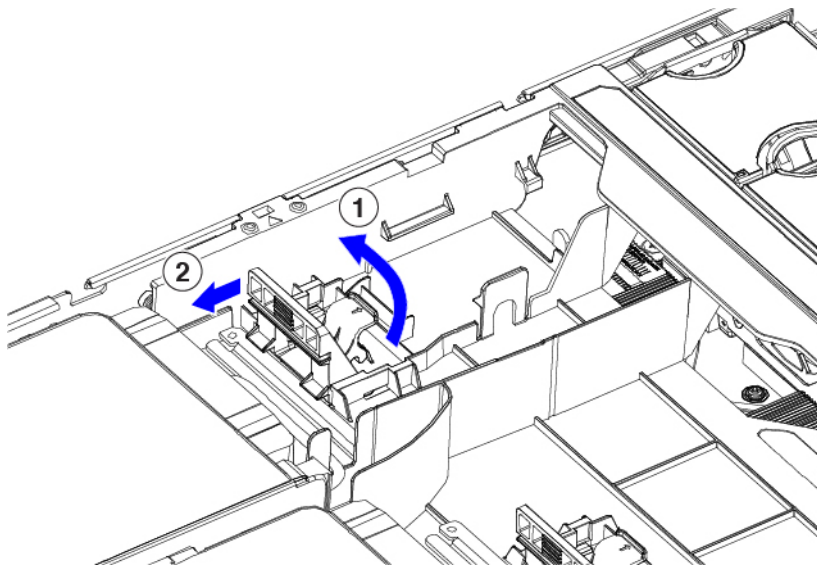
- c) カードの端のカードタブ固定具を閉じます。

- d) ライザーの底部でヒンジ付き固定プレートを回して閉じます。プレートのクリップが、カチッと音がしてロック位置に収まったことを確認します。
- e) PCIe ライザーを、マザーボード上のソケットとシャーシの位置合わせチャネルの上に配置します。
- f) PCIe ライザーの両端を慎重に押し下げて、コネクタをマザーボード上のソケットにしっかりと収納します。

同時に、（GPUカードのフロントエンドにある）GPUの前面支持ブラケットを、サーバのエアージャケットにある固定ラッチの位置に合わせます。

ステップ 6 GPU の前面支持ブラケットをエアージャケット上のラッチに差し込みます。

- a) ラッチリリース タブをつまみ、ラッチをサーバの前面に向け倒します。
- b) ラッチを後方に向けて倒します。これにより、ラッチのへりが、GPU の前面支柱ブラケットの端にかぶるように閉じます。
- c) ラッチリリース タブがカチッと音がしてラッチが所定の位置に固定されたことを確認します。



ステップ 7 上部カバーをサーバに再度取り付けます。

ステップ 8 ラック内のサーバを交換し、電源ケーブルとネットワークケーブルを交換し、電源ボタンを押してサーバの電源を完全にオンにします。

ステップ 9 オプション：GPU カードをサポートするドライバのインストール（24 ページ）に進みます。

ヒートシンクの交換

GPU の場合、正しいヒートシンクはロープロファイルヒートシンク（UCSC-HSLP-C220M7）で、メインヒートシンクに 4 本の T30 トルクス ネジがあり、拡張ヒートシンクに 2 本のプラスネジがあります。高プロファイルヒートシンク（UCSC-HSHP-C240M7）は GPU では使用できません。

GPU のヒートシンクを交換するには、次の手順を使用します。

- ヒートシンクの取り外し (10 ページ)
- ヒートシンクの取り付け (12 ページ)

ヒートシンクの取り外し

この手順により、ロープロファイルヒートシンク (UCSC-HSLP-C220M7) を GPU から取り外します。

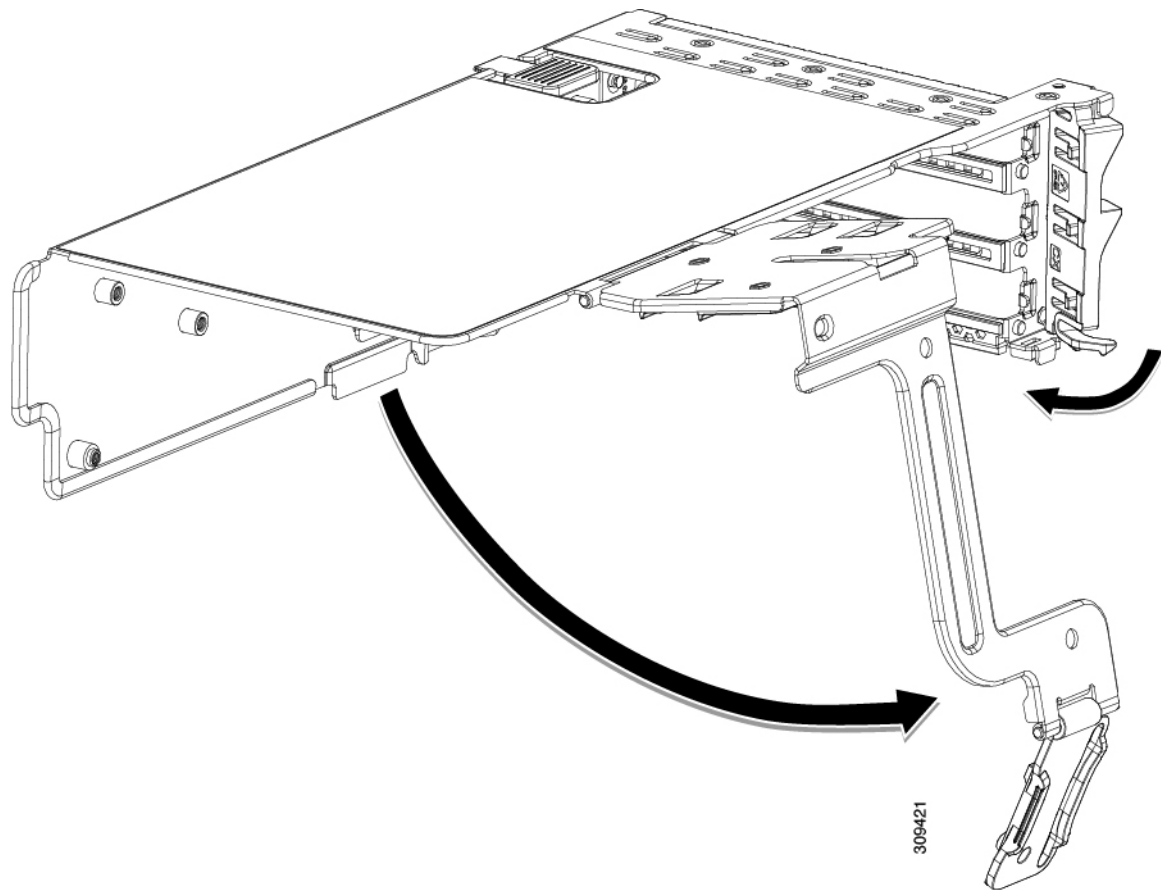
ステップ1 サーバの上部カバーを取り外します。 .

ステップ2 エアダクトを取り外します。 .

ステップ3 ダブル幅の GPU を取り外します。

- a) 両手で PCIe ライザーの金属製ブラケットを持ってまっすぐ持ち上げ、マザーボード上のソケットからコネクタを外します。ライザーを静電気防止素材の上に置きます。
- b) ライザーの底部で、固定プレートを留めているクリップを押し下げます。
- c) ヒンジ付き固定プレートを回して開け、中にアクセスできるようにします。
- d) カードの背面パネル タブを固定しているヒンジ付きプラスチック製固定具を開きます。
- e) PCIe ライザーの電源コネクタから GPU カードの電源ケーブルを外します。
- f) GPU カードの両端を均等に引いて、PCIe ライザーのソケットから取り外します。

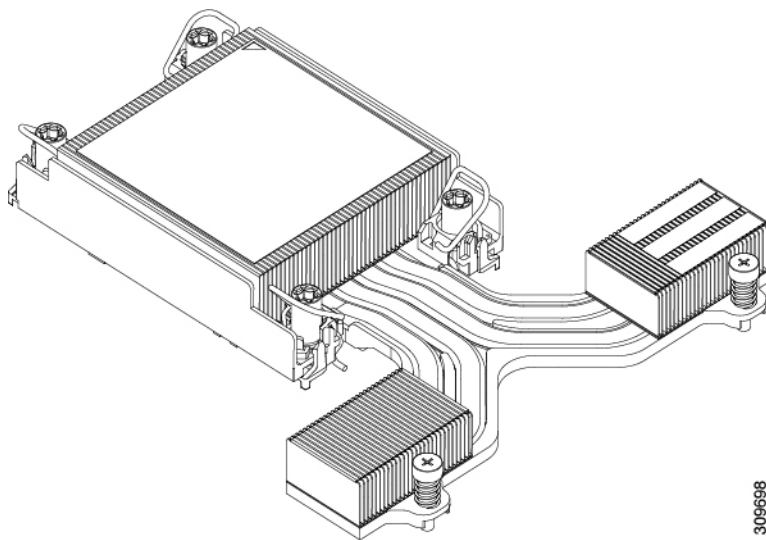
図 2: PCIe ライザーカードの固定機構



1	ヒンジ付き固定プレートのリリース ラッチ	3	ヒンジ付きカード タブ固定具
2	ヒンジ付き固定プレート	-	

ステップ 4 CPU を取り外します。

- a) #2 プラスドライバを使用して、拡張ヒートシンクの 2 本のプラスネジを緩めます。
- b) T30 トルクスドライバを使用して、4 つのトルクス固定ナットを緩めます。



- c) 回転ワイヤを互いに向かって押し、ロック解除位置に移動します。

注意 回転するワイヤができるだけ内側にあることを確認します。完全にロック解除されると、回転するワイヤの下部が外れ、CPU アセンブリを取り外すことができます。回転ワイヤが完全にロック解除位置にない場合、CPU アセンブリを取り外すときに抵抗を感じる場合があります。

- d) キャリアの端に沿って CPU とヒートシンクをつかみ、CPU とヒートシンクを持ち上げてマザーボードから外します。

注意 CPU アセンブリを持ち上げる際は、ヒートシンクフィンを曲げないようにしてください。また、CPU アセンブリを持ち上げるときに抵抗を感じる場合は、回転ワイヤが完全にロック解除位置にあることを確認します。

ステップ 5 GPU からヒートシンクを取り外します。

次のタスク

ロープロファイルヒートシンク (UCSC-HSLP-C220M7) を GPU に取り付けます。「[ヒートシンクの取り付け \(12 ページ\)](#)」を参照してください。

ヒートシンクの取り付け

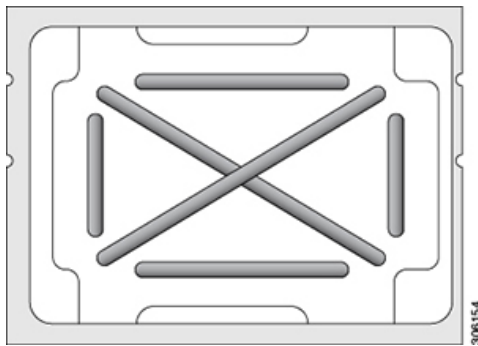
GPU にロープロファイルヒートシンク (UCSC-HSLP-C220M7) を取り付けるには、次の手順を使用します。

ステップ 1 必要に応じて、新しい TIM を適用します。

(注) 適切に冷却し、期待されるパフォーマンスを実現するために、ヒートシンクのCPU側の表面に新しいTIMを塗布する必要があります。

- 新しいヒートシンクを取り付ける場合は、新しいヒートシンクにはTIMが塗布されたパッドが付属しています。ステップ2に進みます。
 - ヒートシンクを再利用する場合は、ヒートシンクから古いTIMを除去してから、付属のシリンジから新しいTIMをCPU表面に塗布する必要があります。次のステップaに進みます。
- a) ヒートシンク クリーニング キット (UCSX-HSCK=) およびスペアの CPU パッケージに同梱されているボトル #1 洗浄液をヒートシンクの古いTIMに塗布し、15秒以上浸しておきます。
 - b) ヒートシンク クリーニング キットに同梱されている柔らかい布を使用して、ヒートシンクからすべてのTIMを拭き取ります。ヒートシンクの表面に傷をつけないように注意してください。
 - c) ボトル #2 を使用してヒートシンクの底面を完全にきれいにして、ヒートシンクの取り付けを準備します。
 - d) 新しいCPU (UCS-CPU-TIM=) に付属のTIMのシリンジを使用して、CPUの上部に1.5立方センチメートル(1.5ml)のサーマルインターフェイス材料を貼り付けます。均一に覆うために、次の図に示すパターンを使用してください。

図3:サーマルインターフェイス材料の貼り付けパターン



注意 正しいヒートシンクのみを使用してください。GPUサーバおよびGPU対応サーバの場合は、UCSC-HSLP-C220M7を使用します。

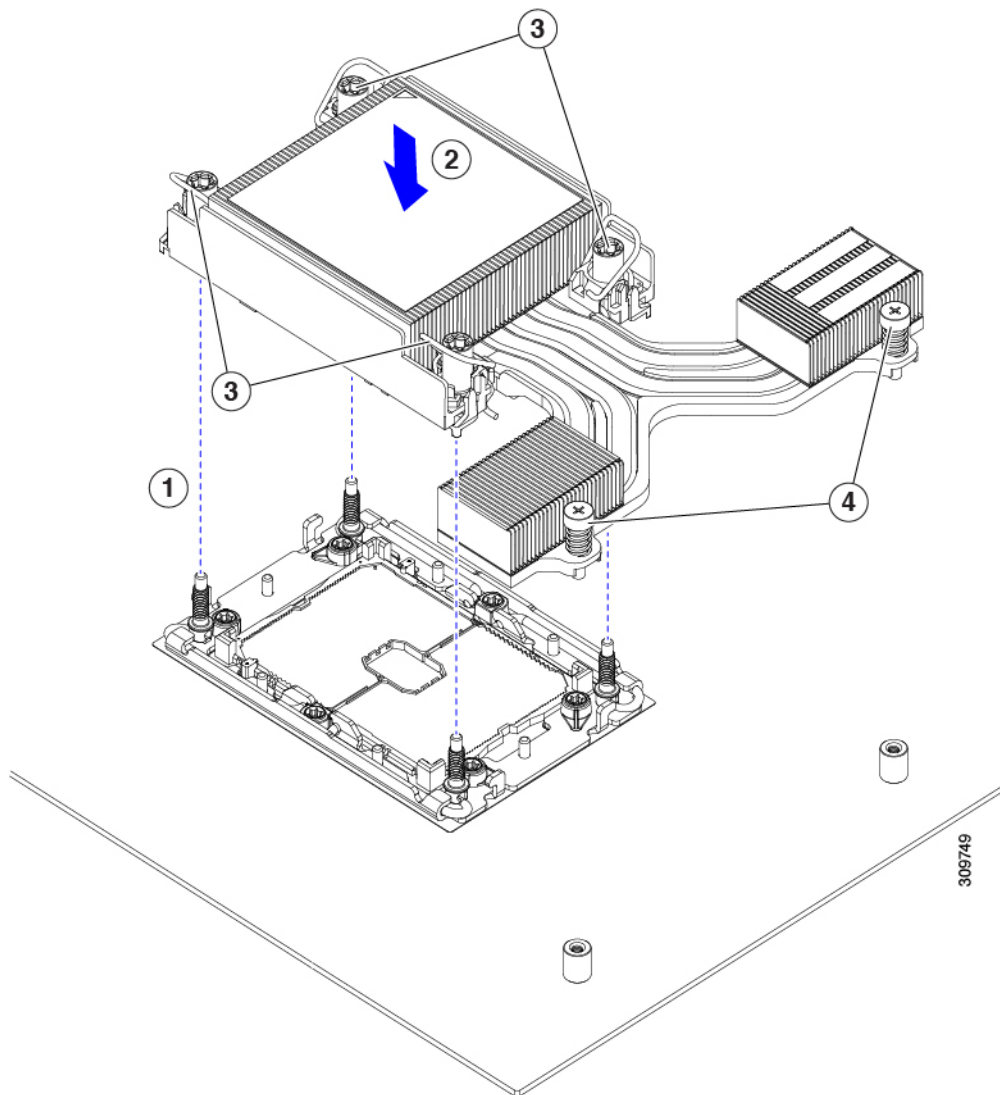
ステップ2 次のように、ヒートシンクを取り付けます。

- a) 取り付けを妨げないように、回転するワイヤをロック解除位置に押しします。
- b) ヒートシンクのフィンをつかみ、ヒートシンクのピン1の位置をGPUのピン1の位置(次の図の2)に合わせ、ヒートシンクをCPUソケットに装着します。
- c) CPUアセンブリを水平に保持し、図のように向きを合わせてCPUソケットの上に置きます。
- d) CPUアセンブリをCPUソケットに固定するために、回転するワイヤを互いに離します。

注意 トルクスドライバを使用して固定ナットを締める前に、回転ワイヤを完全に閉じてください。

- e) T30トルクスドライバを12インチポンドのトルクに設定し、4個の固定ナットを締めてCPUをマザーボードに固定します(3)。

- f) トルクスドライバを6インチポンドのトルクに設定し、拡張ヒートシンク用の2本のプラスネジを締めます (4)。



GPU エア ブロッカーの交換

GPUエアブロッカーは、スロット2、5、または7のPCIeライザーに取り付けられ、上方に伸びてその上のスロットを覆う成形部品です。GPUエアブロッカーは適切なエアフローを提供し、ほこりやその他の潜在的な粒子状物質のレベルを減らします。

GPU構成またはGPU対応サーバの場合、ライザーの一部として必要な場所にGPUエアブロッカーがインストールされます (UCSC-RISAB-24MX7)。エアブロッカーは、別途注文可能な部品 (UCSC-RISAB-24MX7=) としても利用できます。

GPU エア ブロッカーを交換するには、適切なタスクを使用します。



(注) ライザー 1 は GPU エア ブロッカーを受け入れません。サービス タスクを実行する場合は、ライザー 1 のエア ブロッカーを取り付けないでください。

- [ライザー 2 GPU エア ブロッカーを取り外す \(15 ページ\)](#)
- [ライザー 2 GPU エア ブロッカーの取り付け \(16 ページ\)](#)
- [ライザー 3 GPU エア ブロッカーを取り外す \(18 ページ\)](#)
- [ライザー 3 GPU エア ブロッカーの取り付け \(19 ページ\)](#)

ライザー 2 GPU エア ブロッカーを取り外す

GPU エア ブロッカーは、ライザー 2 のスロット 5 に取り付けられ、上方に伸びてその上のスロット (スロット 6) を覆う成形部品です。



注意 GPU 構成または GPU 対応サーバでは、GPU を含まないスロットには GPU エア ブロッカーが必要です！空の GPU スロットがあるままでサーバを動作させないでください。

この手順を使用して、GPU エア ブロッカーを取り外します。

ステップ 1 [サーバのシャットダウンと電源切断](#)の説明に従って、サーバをシャットダウンして電源を切ります。

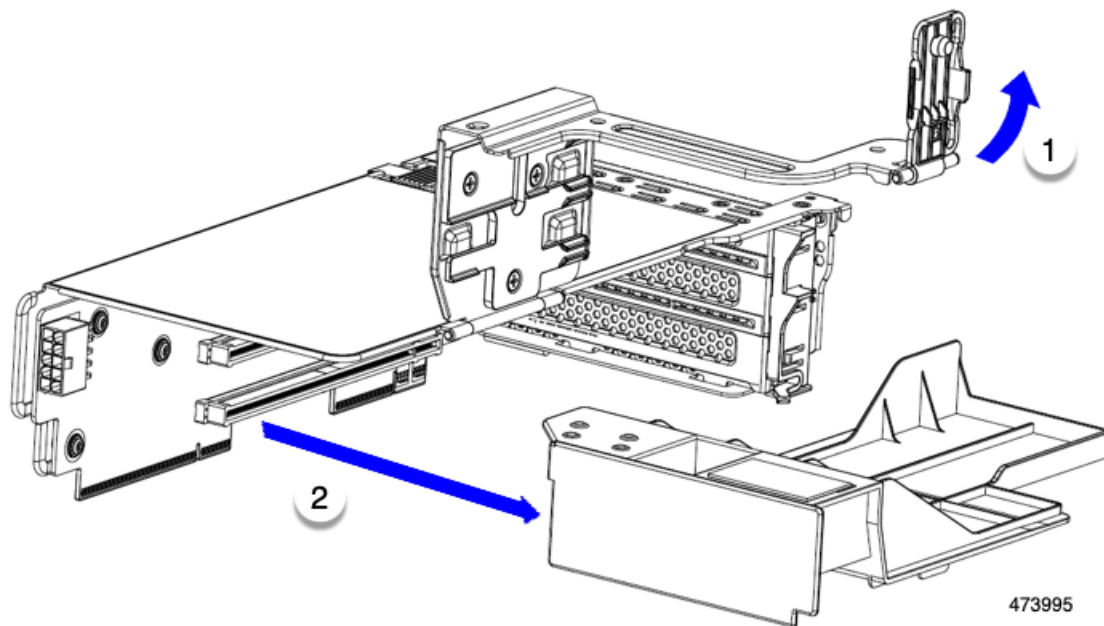
ステップ 2 上部カバーを取り外せるように、サーバをラックの前方に引き出します。場合によっては、リアパネルからケーブルを取り外して隙間を空ける必要があります。

注意 コンポーネントへの接触や確認作業を安全に行えない場合は、ラックからサーバを取り出してください。

ステップ 3 [サーバ上部カバーの取り外し](#)の説明に従ってサーバから上部カバーを取り外します。

ステップ 4 ライザー ケージからエア ブロッカーを取り外します。

- a) 両手で PCIe ライザーの金属製ブラケットを持ってまっすぐ持ち上げ、マザーボード上のソケットからコネクタを外します。
- b) ライザーを静電気防止素材の上に置きます。
- c) ライザーの底部で、固定プレートを留めているクリップを押し下げます。
- d) ヒンジ付き固定プレートを回して開け、中にアクセスできるようにします。
- e) カードの背面パネル タブを固定しているヒンジ付きプラスチック製固定具を開きます。
- f) ブロッカーをつかみ、水平に保ち、ライザー ケージから水平に引き出します。



1	ヒンジ付き固定プレートのリリースラッチ	2	GPU エア ブロッカー
---	---------------------	---	--------------

次のタスク

適切なオプションを選択してください。

- GPU を取り付けます。[倍幅 GPU カードの取り付け \(6 ページ\)](#) を参照してください。
- GPU エア ブロッカーを取り付けます。「[ライザー 2 GPU エア ブロッカーの取り付け \(16 ページ\)](#)」を参照してください。

ライザー 2 GPU エア ブロッカーの取り付け

GPU または GPU エア ブロッカーを取り外した後、GPU またはエア ブロッカーを再取り付けする必要があります。GPU スロットが空いている状態で GPU 構成または GPU 対応サーバを動作させることはできません。

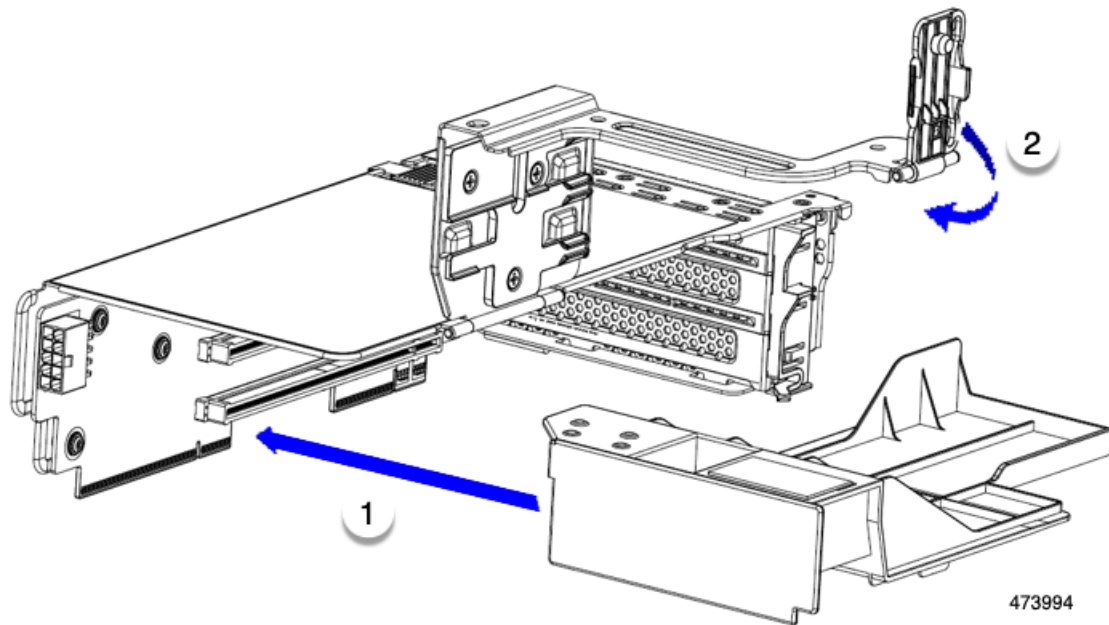
次の手順を使用して、ライザー 2 のスロット 5 に GPU エア ブロッカーを取り付けます。

ステップ 1 スロット 5 のコネクタに挿入するようにパーツの向きを調整します。

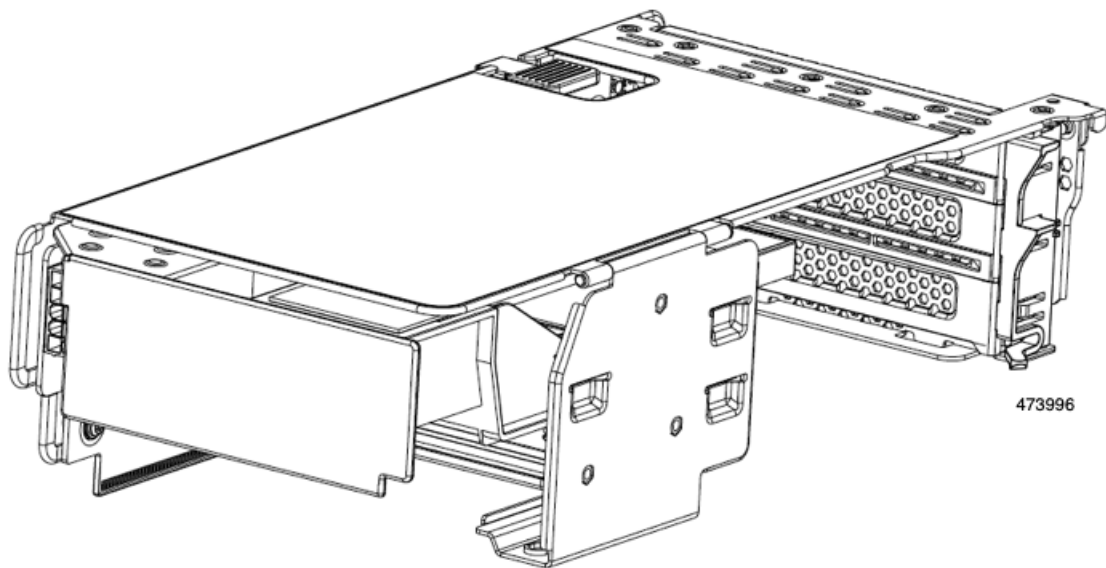
ステップ 2 エア ブロッカーを水平に保ち、ライザー ケージにスライドさせ、ライザー ケージコネクタに挿入します。

エア ブロッカーが取り付けられている場合、エア ブロッカーはライザーにぴったりとはまり、スロット 5 とスロット 6 を覆う必要があります。

ステップ3 ライザー ケージのヒンジ付きドアを閉じてラッチをかけます。



GPU エア ブロッカーが正しく取り付けられると、ヒンジ付きドアが簡単に閉じ、エア ブロッカーが水平になり、ライザー ケージにしっかりと収まります。



ステップ4 ライザー ケージをサーバに取り付けます。

次のタスク

その他の保守作業が必要ない場合は、上部カバーを元に戻し、サーバを稼働させます。それ以外の場合は、追加のメンテナンス タスクを続行します。

ライザー 3 GPU エア ブロッカーを取り外す

GPU エア ブロッカーは、ライザー 3 のスロット 7 に取り付けられ、上方に伸びてその上のスロット (スロット 8) を覆う成形部品です。



注意 GPU 構成または GPU 対応サーバでは、GPU を含まないスロットには GPU エア ブロッカーが必要です！空の GPU スロットがあるままでサーバを動作させないでください。

この手順を使用して、GPU エア ブロッカーを取り外します。

ステップ 1 サーバのシャットダウンと電源切断の説明に従って、サーバをシャットダウンして電源を切ります。

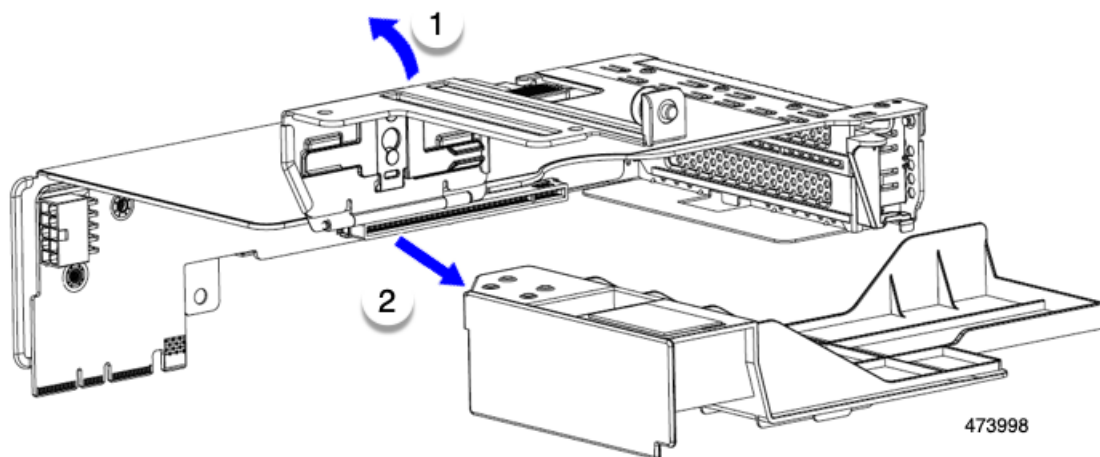
ステップ 2 上部カバーを取り外せるように、サーバをラックの前方に引き出します。場合によっては、リアパネルからケーブルを取り外して隙間を空ける必要があります。

注意 コンポーネントへの接触や確認作業を安全に行えない場合は、ラックからサーバを取り出してください。

ステップ 3 サーバ上部カバーの取り外しの説明に従ってサーバから上部カバーを取り外します。

ステップ 4 ライザー ケージからエア ブロッカーを取り外します。

- a) 両手で PCIe ライザーの金属製ブラケットを持ってまっすぐ持ち上げ、マザーボード上のソケットからコネクタを外します。
- b) ライザーを静電気防止素材の上に置きます。
- c) ライザーの底部で、固定プレートを留めているクリップを押し下げます。
- d) ヒンジ付き固定プレートを回して開け、中にアクセスできるようにします。
- e) カードの背面パネル タブを固定しているヒンジ付きプラスチック製固定具を開きます。
- f) ブロッカーをつかみ、水平に保ち、ライザー ケージから水平に引き出します。



1	ヒンジ付き固定プレートのリリースラッチ	2	GPU エア ブロッカー
---	---------------------	---	--------------

次のタスク

適切なオプションを選択してください。

- GPU を取り付けます。倍幅 GPU カードの取り付け (6 ページ) を参照してください。
- GPU エア ブロッカーを取り付けます。「ライザー 3 GPU エア ブロッカーの取り付け (19 ページ)」を参照してください。

ライザー 3 GPU エア ブロッカーの取り付け

GPU または GPU エア ブロッカーを取り外した後、GPU またはエア ブロッカーを再取り付けする必要があります。GPU スロットが空いている状態で GPU 構成または GPU 対応サーバを動作させることはできません。

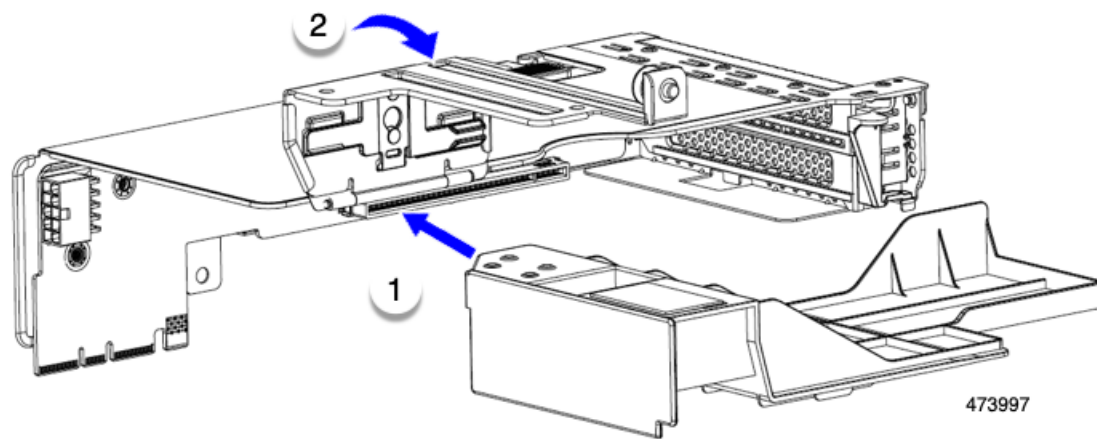
次の手順を使用して、ライザー 3 のスロット 7 に GPU エア ブロッカーを取り付けます。

ステップ 1 スロット 7 のライザー ケージ コネクタに挿入されるように部品の向きを合わせます。

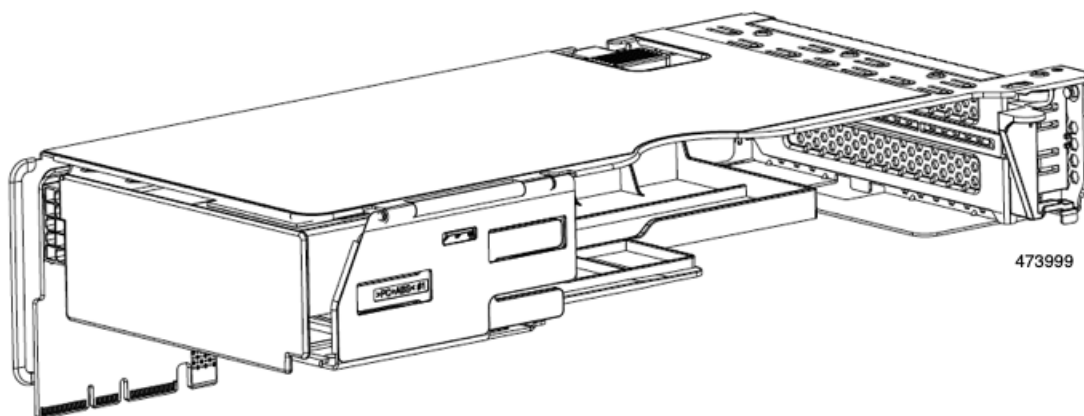
ステップ 2 エア ブロッカーを水平に保ち、ライザー ケージにスライドさせます。

エア ブロッカーが取り付けられている場合、エア ブロッカーはライザーにぴったりとはまり、スロット 7 とスロット 8 を覆う必要があります。

ステップ 3 ライザー ケージのヒンジ付きドアを閉じます。



GPU エア ブロッカーが正しく取り付けられると、ヒンジ付きドアが簡単に閉じ、エア ブロッカーが水平になり、ライザー ケージにしっかりと収まります。



ステップ 4 ライザー ケージをサーバに取り付けます。

次のタスク

その他の保守作業が必要ない場合は、上部カバーを元に戻し、サーバを稼働させます。それ以外の場合は、追加のメンテナンス タスクを続行します。

NVIDIA ライセンス ポータルからライセンス サーバへの グリッドライセンスのインストール

グリッドライセンス サーバ管理インターフェイスへのアクセス

ライセンス サーバ ホストの Web ブラウザを開き、URL <http://localhost:8080/licserver> にアクセスします。

ライセンス サーバへのリモートアクセスを許可するようにライセンス サーバホストのファイアウォールを設定した場合は、管理インターフェイスに <http://hostname:8080/licserver> の URL でリモート マシンからアクセスできます。

ライセンス サーバの MAC アドレスの読み取り

ライセンス サーバのイーサネット MAC アドレスは NVIDIA のライセンス ポータルでライセンス サーバを登録するときに ID として使用されます。

ステップ 1 ブラウザで GRID ライセンス サーバ管理インターフェイスにアクセスします。

ステップ 2 左側の [ライセンス サーバ (License Server)] パネルで [設定 (Configuration)] を選択します。

[ライセンス サーバの設定 (License Server Configuration)] パネルが開きます。[サーバのホスト ID (Server host ID)] の横のプルダウンメニューに、選択可能なイーサネット MAC アドレスがリストされます。

ステップ 3 [サーバのホスト ID (Server host ID)] プルダウンからライセンス サーバの MAC アドレスを選択します。

(注) NVIDIA のライセンス ポータルでライセンスを生成する場合には、サーバを識別するために一貫して同じイーサネット ID を使用することが重要です。NVIDIA では、プラットフォーム上の、削除できないプライマリーイーサネットインターフェイス用にエントリを1つ選択することを推奨しています。

ライセンス ポータルからのライセンスのインストール

ステップ 1 ブラウザで GRID ライセンス サーバ管理インターフェイスにアクセスします。

ステップ 2 左側の [ライセンス サーバ (License Server)] パネルで [設定 (Configuration)] を選択します。

[ライセンス サーバの設定 (License Server Configuration)] パネルが開きます。

ステップ 3 前に生成した .bin ファイルをインストールするには、[License Server Configuration] メニューを使用します。

- [ファイルを選択 (Choose File)] をクリックします。
- インストールするライセンス .bin ファイルを参照して、[Open] をクリックします。
- [アップロード (Upload)] をクリックします。

ライセンス サーバにライセンス ファイルがインストールされます。インストールが完了すると、「Successfully applied license file to license server」という確認メッセージが表示されます。

使用可能なグリッドライセンスの表示

インストールされて利用可能なライセンスとそのプロパティを表示するには、次の手順を使用します。

-
- ステップ1 ブラウザで GRID ライセンス サーバ管理インターフェイスにアクセスします。
 - ステップ2 左側の [ライセンス サーバ (License Server)] パネルで [ライセンス機能の使用 (Licensed Feature Usage)] を選択します。
 - ステップ3 [機能 (Feature)] 列の機能をクリックすると、その機能の現在の使用に関する詳細情報が表示されます。
-

現在のライセンスの使用状況の表示

現在使用中であり、サーバから交付されているライセンスに関する情報を表示するには、次の手順を実行します。

-
- ステップ1 ブラウザで GRID ライセンス サーバ管理インターフェイスにアクセスします。
 - ステップ2 左側の [ライセンス サーバ (License Server)] パネルで [ライセンスされたクライアント (Licensed Clients)] を選択します。
 - ステップ3 シングルライセンス クライアントに関する詳細情報を表示するには、リストの [クライアント ID (Client ID)] をクリックします。
-

グリッドライセンスの管理

グリッドライセンスを必要とする機能は、グリッドライセンスを取得するまで、機能が限定された状態で実行されます。

Windows での GRID ライセンスの取得

-
- ステップ1 次のいずれかの方法を使用して NVIDIA コントロール パネルを開きます。
 - Windows デスクトップを右クリックして、メニューから [NVIDIA Control Panel] を選択します。
 - Windows コントロール パネルを開き、[NVIDIA Control Panel] アイコンをダブルクリックします。
 - ステップ2 NVIDIA コントロール パネルの左側のペインで、[Licensing] の下の [Manage License] を選択します。

[Manage License] タスク ペインが開き、現在使用されているライセンス エディションが表示されます。グリッドソフトウェアは、使用している機能に基づいてライセンス エディションを自動的に選択します。デフォルトは、Tesla (ライセンスなし) です。

- ステップ3** グリッド仮想ワークステーションのライセンスを取得するには、[License Edition] で **[GRID Virtual Workstation]** を選択します。
- ステップ4** **[ライセンス サーバ (License Server)]** フィールドに、ローカルの GRID ライセンス サーバのアドレスを入力します。アドレスには、ドメイン名または IP アドレスを指定できます。
- ステップ5** **[Port Number]** フィールドに、サーバが使用するポート番号を入力するか、またはデフォルト設定 (7070) のままにしておきます。
- ステップ6** **[適用 (Apply)]** を選択します。

システムは、設定されているライセンスサーバから適切なライセンスエディションを要求します。ライセンスが正常に取得されると、そのライセンス エディションの機能が有効になります。

(注) **[NVIDIA Control Panel (NVIDIA コントロール パネル)]** でライセンスを設定すると、その設定はリブート後も保持されます。

Linux での GRID ライセンスの取得

- ステップ1** コンフィギュレーション ファイル `/etc/nvidia/gridd.conf` を編集します。

```
sudo vi /etc/nvidia/gridd.conf
```

- ステップ2** ローカル グリッド ライセンス サーバのアドレスを使用して `ServerUrl` の行を編集します。
アドレスには、ドメイン名または IP アドレスを指定できます。次のファイルの例を参照してください。
- ステップ3** コロンを使用してアドレスの最後にポート番号 (デフォルトは 7070) を追加します。次のファイルの例を参照してください。
- ステップ4** ライセンス タイプの整数を使用して `FeatureType` の行を編集します。次のファイルの例を参照してください。
- グリッド vGPU = 1
 - グリッド仮想ワークステーション = 2

- ステップ5** `nvidia-gridd` サービスを再起動します。

```
sudo service nvidia-gridd restart
```

サービスは自動的に、`FeatureType` 行に指定したライセンス エディションを取得します。これは、`/var/log/messages` で確認できます。

(注) **NVIDIA コントロール パネル** でライセンスを設定すると、その設定はリブート後も保持されません。

サンプル コンフィギュレーション ファイル :

```
# /etc/nvidia/gridd.conf - Configuration file for NVIDIA Grid Daemon
# Description: Set License Server URL
# Data type: string
```

```
# Format: "<address>:<port>"
ServerUrl=10.31.20.45:7070

# Description: Set Feature to be enabled
# Data type: integer
# Possible values:
# 1 => for GRID vGPU
# 2 => for GRID Virtual Workstation
FeatureType=2
```

gpumodeswitch の使用

コマンドラインユーティリティ `gpumodeswitch` は、次の環境で実行できます。

- Windows 64 ビットのコマンド プロンプト（管理者権限が必要）
- Linux 32/64 ビット シェル（Citrix XenServer dom0 を含む）（ルート権限が必要）



(注) コンピューティング モードおよびグラフィック モードとの互換性の最新情報については、NVIDIA 製品のリリース ノートを参照してください。

`gpumodeswitch` ユーティリティでは、次のコマンドがサポートされています。

- `-listgpumodes`

このコマンドは、現在の作業ディレクトリにある `listgpumodes.txt` というログ ファイルに情報を書き込みます。

- `--gpumode graphics`

グラフィック モードに切り替えます。プロンプトが表示された際に、特別に指定しない限り、サーバでサポートされているすべての GPU のモードを切り替えます。

- `--gpumode compute`

コンピューティングモードに切り替えます。プロンプトが表示された際に、特別に指定しない限り、サーバでサポートされているすべての GPU のモードを切り替えます。



(注) GPU モードを切り替えた後、サーバを再起動して、GPU の修正したリソースがサーバで実行されている OS またはハイパーバイザによって正しく認識されることを確認してください。

GPU カードをサポートするドライバのインストール

ハードウェアの取り付け後、サーバ BIOS を適切なレベルに更新し、ドライバなどのソフトウェアを次の順序でインストールする必要があります。

1. サーバ BIOS を更新します。
2. GPU ドライバを更新します。

1. サーバ BIOS の更新

Host Upgrade Utility を使用して、最新の Cisco UCS C240 M4 サーバ BIOS を Cisco UCS C240 M4 サーバにインストールします。



(注) NVIDIA ドライバを更新する前に、次の手順を実行する必要があります。

- ステップ 1 <http://www.cisco.com/cisco/software/navigator.html> にアクセスします。
- ステップ 2 中央の列の [サーバ - ユニファイドコンピューティング (Servers - Unified Computing)] をクリックします。
- ステップ 3 右側の列の [Cisco UCS C シリーズラックマウントスタンドアロンサーバソフトウェア (UCS C-Series Rack-Mount Standalone Server Software)] をクリックします。
- ステップ 4 右側のカラムでお使いのサーバのモデルの名前をクリックします。
- ステップ 5 [Unified Computing System (UCS) サーバソフトウェア (Unified Computing System (UCS) Server Firmware)] をクリックします。
- ステップ 6 リリース番号をクリックします。
- ステップ 7 [今すぐダウンロード (Download Now)] をクリックして `ucs-server platform-huu-version_number.iso` ファイルをダウンロードします。
- ステップ 8 次のページで情報を確認した後、[ダウンロードを続行する (Proceed With Download)] をクリックします。
- ステップ 9 次の画面に進んでライセンス契約に同意し、このファイルを保存する場所を参照します。
- ステップ 10 サーバ BIOS を更新するには、Host Upgrade Utility を使用します。

Host Upgrade Utility のユーザ ガイドは、『[Utility User Guides](#)』を参照してください。

2. GPU カード ドライバの更新

サーバ BIOS を更新したら、ハイパーバイザ仮想マシンに GPU ドライバをインストールできます。

- ステップ 1 コンピュータにハイパーバイザソフトウェアをインストールします。インストール手順については、ハイパーバイザのマニュアルを参照してください。
- ステップ 2 ハイパーバイザ内で仮想マシンを作成します。手順については、ハイパーバイザのマニュアルを参照してください。

ステップ3 仮想マシンにGPUドライバをインストールします。ドライバを次のいずれかのサイトからダウンロードします。

- NVIDIA エンタープライズ ポータル、GRID ハイパーバイザ ダウンロード (NVIDIA ログインが必要です) : <https://nvidia.flexnetoperations.com/>
- NVIDIA パブリック ドライバ エリア : <http://www.nvidia.com/Download/index.aspx>

ステップ4 サーバを再起動します。

ステップ5 仮想マシンがGPUカードを認識できることを確認します。Windowsでは、[デバイスマネージャー (Device Manager)]の[ディスプレイアダプター (Display Adapters)]から確認します。

翻訳について

このドキュメントは、米国シスコ発行ドキュメントの参考和訳です。リンク情報につきましては、日本語版掲載時点で、英語版にアップデートがあり、リンク先のページが移動/変更されている場合がありますことをご了承ください。あくまでも参考和訳となりますので、正式な内容については米国サイトのドキュメントを参照ください。