



はじめに

- [仮想化の概要 \(1 ページ\)](#)
- [Cisco Virtual Machine ファブリック エクステンダの概要 \(1 ページ\)](#)
- [仮想インターフェイス カード アダプタでの仮想化 \(2 ページ\)](#)
- [単一のルート I/O 仮想化 \(2 ページ\)](#)
- [KVM 用 VM-FEX \(4 ページ\)](#)

仮想化の概要

仮想化により、独立して実行する複数の仮想マシン (VM) を同一の物理マシン上に隣接させて作成できます。

各仮想マシンは、仮想ハードウェア (メモリ、CPU、NIC) の独自のセットを持ち、その上でオペレーティングシステムと十分に設定されたアプリケーションがロードされます。オペレーティングシステムは、実際の物理ハードウェア コンポーネントに関係なく、一貫性があり正常なハードウェア一式を認識します。

仮想マシンでは、物理サーバ間でのプロビジョニングや移動を迅速に行うために、ハードウェアとソフトウェアの両方が単一のファイルにカプセル化されます。仮想マシンは1つの物理サーバから別のサーバへ数秒で移動することができ、メンテナンスのためのダウンタイムを必要とせず、途切れることのない作業負荷を集約します。

仮想ハードウェアは、多数のサーバ (それぞれのサーバは独立した仮想マシン内で実行する) を単一の物理サーバ上で実行できるようにします。仮想化の利点は、コンピューティングリソースをより適切に使用でき、サーバ密度を高め、サーバの移行をスムーズに行えることです。

Cisco Virtual Machine ファブリック エクステンダの概要

仮想サーバの実装は、1つの物理サーバのゲストとして実行される1つまたは複数の VM で構成されます。ゲスト VM は、ハイパーバイザまたは仮想マシンマネージャ (VMM) と呼ばれるソフトウェアレイヤによってホストされ管理されます。通常、ハイパーバイザは各 VM への仮想ネットワークインターフェイスを示し、VM から他のローカル VM へのトラフィックの

レイヤ2スイッチング、または外部ネットワークに対する別のインターフェイスへのトラフィックのレイヤ2スイッチングを実行します。

Cisco 仮想インターフェイス カード (VIC) アダプタと連携して、Cisco Virtual Machine ファブリック エクステンダ (VM-FEX) はファブリック インターコネクットの外部ハードウェア ベース スwitching用のハイパーバイザによって、VM トラフィックのソフトウェア ベースのスイッチングをバイパスします。この方法により、サーバの CPU 負荷を軽減し、高速スイッチングを行い、ローカルおよびリモートトラフィックに豊富なネットワーク管理機能セットを適用することができます。

VM-FEX は IEEE 802.1Qbh ポート エクステンダ アーキテクチャを VM に拡張するために、各 VM インターフェイスに仮想 Peripheral Component Interconnect Express (PCIe) デバイスとスイッチ上の仮想ポートを提供します。このソリューションにより、VM インターフェイス上で、正確なレート制限と QoS (Quality of Service) 保証が可能になります。



重要 Cisco UCS Manager リリース 4.0(1) では、VM-FEX は Cisco UCS 6454 ファブリック インターコネクットではサポートされていません。

仮想インターフェイス カード アダプタでの仮想化

Cisco VIC アダプタは、ベア メタルの導入と VM ベースの導入の両方に対応するように設計された、統合型ネットワーク アダプタ (CNA) です。VIC アダプタは、最大 116 個の仮想ネットワーク インターフェイス カード (vNIC) を含む、静的または動的な仮想化インターフェイスをサポートします。

VIC アダプタに使用される vNICs には、静的と動的の 2 つのタイプがあります。静的な vNIC は、OS またはハイパーバイザから認識されるデバイスです。動的な vNIC は、VM をファブリック インターコネクットの vEth ポートに接続するための VM-FEX に使用されます。

VIC アダプタは、VM-FEX をサポートし、仮想マシンインターフェイスとの間の、トラフィックのハードウェアベースのスイッチング機能を提供します。

単一のルート I/O 仮想化

Single Root I/O Virtualization (SR-IOV) により、さまざまなゲスト オペレーティング システムを実行している複数の VM が、ホストサーバ内の単一の PCIe ネットワーク アダプタを共有できるようになります。SR-IOV では、VM がネットワーク アダプタとの間で直接データを移動でき、ハイパーバイザをバイパスすることで、ネットワークのスループットが増加しサーバの CPU 負荷が低下します。最近の x86 サーバプロセッサには、SR-IOV に必要なダイレクトメモリの転送やその他の操作を容易にする Intel VT x テクノロジーなど、チップセットの拡張機能が搭載されています。

SR-IOV 仕様では、次の 2 つのデバイス タイプが定義されています。

- **Physical Function (PF)** : 基本的にスタティック vNIC です。PF は、SR-IOV 機能を含む完全な PCIe デバイスです。PF は、通常の PCIe デバイスとして検出、管理、設定されます。単一 PF は、一連の仮想関数 (VF) の管理および設定を提供できます。
- **Virtual Function (VF)** : ダイナミック vNIC に似ています。VF は、データ移動に必要な最低限のリソースを提供する、完全または軽量の仮想 PCIe デバイスです。VF は直接的には管理されず、PF を介して配信および管理されます。1 つ以上の VF を 1 つの VM に割り当てることができます。

SR-IOV は、PCI 標準の開発および管理が公認されている業界組織である Peripheral Component Interconnect Special Interest Group (PCI-SIG) によって定義および管理されています。SR-IOV の詳細については、次の URL を参照してください。

<https://www.intel.com/content/www/us/en/pci-express/pci-sig-sr-iov-primer-sr-iov-technology-paper.html>

Linux KVM ハイパーバイザでは、SR IOV がサポートされています。

次のシスコ仮想インターフェイスカードは VM-FEX を使用する SR-IOV をサポートしていません。

- Cisco UCS 仮想インターフェイスカード 1240
- Cisco UCS 仮想インターフェイスカード 1280
- Cisco UCS 仮想インターフェイスカード 1225
- Cisco UCS 仮想インターフェイスカード 1225T
- Cisco UCS 仮想インターフェイスカード 1227
- Cisco UCS 仮想インターフェイスカード 1227T
- Cisco UCS 仮想インターフェイスカード 1340
- Cisco UCS 仮想インターフェイスカード 1380
- Cisco UCS 仮想インターフェイスカード 1385
- Cisco UCS 仮想インターフェイスカード 1387
- Cisco UCS 仮想インターフェイスカード 1440
- Cisco UCS 仮想インターフェイスカード 1480
- Cisco UCS 仮想インターフェイスカード 1455
- Cisco UCS 仮想インターフェイスカード 1457

KVM 用 VM-FEX

KVM 用 VM-FEX の概要

カーネルベースの仮想マシン（KVM）は、x86 ハードウェアプラットフォームの Linux 向け仮想化パッケージです。KVM は x86 ハードウェア仮想化拡張機能（たとえば Intel VT-x）を使用して、VM をホストするハイパーバイザを、ユーザ空間プロセスとして実装します。

KVM 用の VM-FEX を使用すると、ハイパーバイザは VM トラフィックのスイッチングを行いません。インストールされている VIC アダプタを使用して、ハイパーバイザはインターフェイスバーチャライザとして機能し、次の機能を実行します。

- VM から VIC へのトラフィックについては、vNIC によって生成される各パケットに VIC が明示的にタグ付けできるように、インターフェイスバーチャライザが発信元の vNIC を識別します。
- VIC から受信したトラフィックの場合、インターフェイスバーチャライザは指定された vNIC にパケットを送信します。

すべてのスイッチングは外部ファブリックインターコネクタによって実行されます。外部ファブリックインターコネクタは、物理ポート間のみでスイッチングを行うのではなく、VM 上の vNIC に対応する仮想インターフェイス（VIF）間でもスイッチングを行います。

KVM の詳細については、次の URL を参照してください。 <https://www.linux-kvm.org>。

Cisco UCS Manager のコンポーネント

クラスタ

Cisco UCS クラスタは、複数のホストにわたって配布できるハイパーバイザをグループ化したものです。KVM システムでは、クラスタは VMware ESX システムの分散仮想スイッチ（DVS）にほぼ対応します。

現在の Cisco UCS KVM 実装では、クラスタはポートプロファイルの範囲を定義し、移行ドメインの境界になっています。複数の KVM ホストがクラスタに関連付けられている場合は、KVM ホストからクラスタ内の残りの部分に VM を移行できます。



- (注) KVM 用 VM-FEX の現在の Cisco UCS の実装では、1つのクラスタ、つまりデフォルトクラスタのみが使用されます。追加のクラスタを作成できますが、KVM ホスト上の VM に対してデフォルトクラスタのみを指定できます。

ポート プロファイル

ポート プロファイルには、Cisco UCS 仮想インターフェイスを設定するために使用する設定とプロパティが含まれています。ポート プロファイルは、Cisco UCS Manager で作成および管理されます。



重要

クラスタによってポートプロファイルが作成され、割り当てられ、アクティブに使用された後に、Cisco UCS Manager でポート プロファイルのネットワーキングプロパティが変更されると、変更がすぐにそれらのクラスタに適用され、ホストのリブートは必要ありません。

Cisco UCS 6454 ファブリック インターコネクトは、ポート プロファイルおよび分散型仮想スイッチ (DVS) に関連する設定をサポートしません。

ポート プロファイル クライアント

ポート プロファイル クライアントは、ポート プロファイルが適用されるクラスタです。



- (注) KVM 用 VM-FEX の現在の Cisco UCS の実装では、デフォルト クラスタが使用できる唯一のポート プロファイル クライアントです。

KVM のコンポーネント

ハイパーバイザ

ハイパーバイザは、VM とネットワーク間の接続により、さまざまなゲストオペレーティングシステムを実行する複数の VM をサポートします。KVM のハイパーバイザは、Red Hat Enterprise Linux (RHEL) がインストールされたホスト サーバです。VM-FEX の最も早いサポート対象リリースは RHEL 6.1 ですが、一部の機能 (SR-IOV など) にはそれより後のバージョンが必要です。

ハイパーバイザを持つホスト サーバには、Cisco VIC アダプタがインストールされている必要があります。

Red Hat Enterprise Linux を使用した仮想化の詳細については、URL <https://www.redhat.com> にある『Red Hat Enterprise Virtualization for Servers Installation Guide』を参照してください。

libvirt

libvirt は、KVM、Xen、VMware ESX などのさまざまな仮想化テクノロジーを管理できるオープン ソース ツールキットです。libvirtd という名前のサービスとしてハイパーバイザで稼働する Libvirt は、コマンドラインインターフェイス (virsh) を提供し、グラフィカルユーザインターフェイスのパッケージ (virt-manager) のツールキットを提供します。

libvirt によって作成および管理される各仮想マシンは、ドメイン XML ファイルの形で表現されます。

libvirt 仮想化 API の詳細については、次の URL を参照してください。 <https://www.libvirt.org>。

virsh CLI の詳細については、次の URL を参照してください。

- <https://linux.die.net/man/1/virsh>
- <https://www.libvirt.org/virshcmdref.html>

MacVTap

MacVTap は、VM の NIC をホスト サーバ上の物理 NIC に直接接続できるようにするための Linux ドライバです。

MacVTap ドライバの詳細については、次の URL を参照してください。 <https://virt.kernelnewbies.org/MacVTap>。

VirtIO

VirtIO のパラ仮想化されたネットワーク ドライバ (`virtio-net`) は、VM のゲストオペレーティング システムで動作し、VM に仮想化認識でエミュレートされたネットワーク インターフェイスを提供します。

VirtIO ドライバの詳細については、次の URL を参照してください。 <https://wiki.libvirt.org/page/Virtio>。

ドライバトポロジ

ドライバのトポロジ (モード) を使用して、VM の vNIC とホストの VIC アダプタ間の VM-FEX 接続を設定できます。これらの各トポロジで、VM トラフィックは VIC アダプタとの間でのみ送受信されます。同じホスト上のある VM から別の VM へのトラフィックは、まず外部ファブリック インターコネクトによるスイッチングのためにホストから出て行く必要があります。



-
- (注) すべてのトポロジにおいて、クイックエミュレータ (QEMU) PCI 層の設定により、ホストが VM に割り当てることができる PCI デバイス数が制限される場合があります。
-

MacVTap Direct (プライベート)

MacVTap Linux ドライバは、ハイパーバイザ (VMM) にインストールされ、各 VM の VirtIO インターフェイスを VIC アダプタの物理 PCIe ポートに接続します。MacVTap ドライバのモードは「プライベート」であり、これは外部スイッチングを使用してすべての VM トラフィックをホスト アダプタとの間で直接送受信することを指定します。サポートされる VM の数は、VIC アダプタ ポートの数に制限されます。ライブ移行がサポートされています。

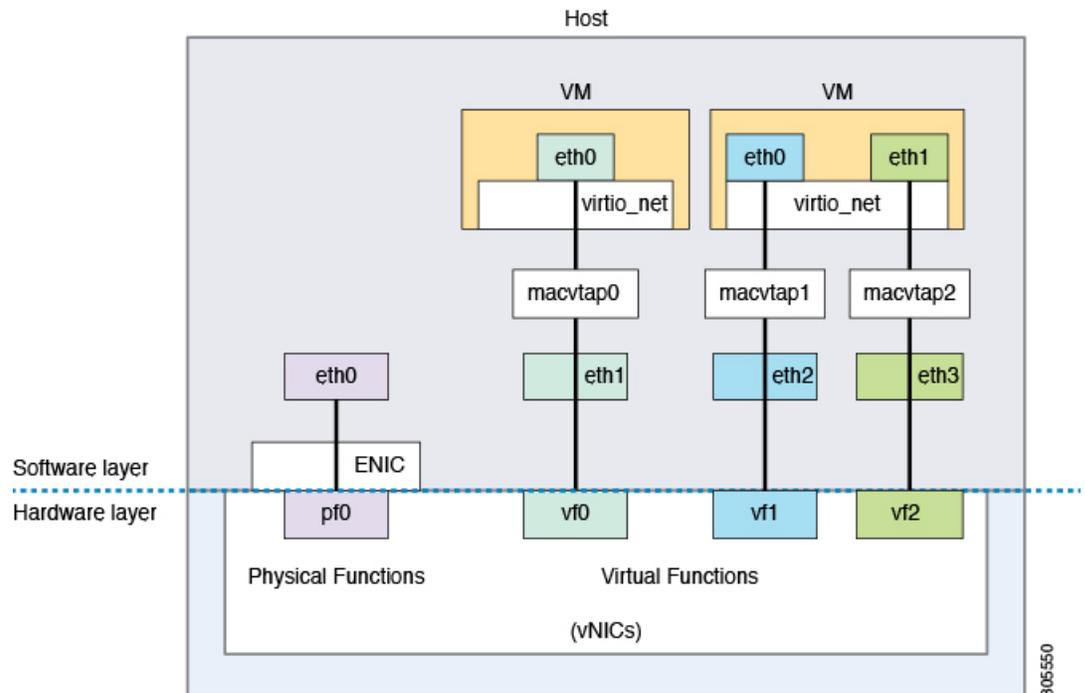


-
- (注) Cisco UCS Release 2.1 以降では、MacVTap の直接 (プライベート) トポロジはサポートされなくなりました。
-

MacVTap パススルー（エミュレーションモード）を使用する SR-IOV

MacVTap の Linux ドライバは、ハイパーバイザにインストールされ、各 VM の VirtIO インターフェイスを SR-IOV 対応 VIC アダプタの VF に接続します。MacVTap ドライバモードは「パススルー」であり、すべての VM トラフィックは VF との間で送受信されます。VF にポートプロファイルを適用する場合、libvirt は VF と関連する PF を決定し、PF を通過する VF を設定します。このトポロジは、MacVTap パススルー（エミュレーションモード）とも呼ばれています。MacVTap パススルーを使用する SR-IOV の例を図 1 に示します。図 1 は、ハードウェアおよびソフトウェアコンポーネントの単純化されたバージョンです。

図 1: 図 1



VM の最大サポート数は、VIC アダプタにおける VF の数によって決まります。PF に割り当てることができる VF の数は、ホストの Netlink プロトコルの実装によってさらに制限される可能性があります（通常は PF あたり 22 ~ 32 VF、OS バージョンごとに異なる）。ライブ移行がサポートされています。

SR-IOV VF パススルー（Hostdev モード）

MacVTap ドライバと VirtIO ドライバは使用されません。代わりに、VIC アダプタのイーサネットドライバ (enic) が VM カーネルにインストールされ、VF に直接接続します。libvirt を使用して、関連する PF を介して VF を設定できます。libvirt のマニュアルで、このトポロジは hostdev モードと呼ばれます。このトポロジは、PCI パススルーとも呼ばれています。サポートされる VM の数は、VIC アダプタで提供されている VF の数で決まります。ライブ移行はサポートされません。VF パススルーを使用する SR-IOV の例を図 2 に示します。図 2 は、ハードウェアおよびソフトウェアコンポーネントの単純化されたバージョンです。

図 2: 図 2

