



# Red Hat Enterprise Virtualization 3.6

## SR-IOV 実装に関するハードウェアの考慮事項

SR-IOV を Red Hat Enterprise Virtualization に実装するにあたってのハードウェア  
の考慮事項



# Red Hat Enterprise Virtualization 3.6 SR-IOV 実装に関するハードウェア の考慮事項

---

SR-IOV を Red Hat Enterprise Virtualization に実装するにあたってのハードウェアの考慮事項

Red Hat Enterprise Virtualization Documentation Team

Red Hat Customer Content Services

[rhev-docs@redhat.com](mailto:rhev-docs@redhat.com)

## 法律上の通知

Copyright © 2018 Red Hat, Inc.

The text of and illustrations in this document are licensed by Red Hat under a Creative Commons Attribution-Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at

<http://creativecommons.org/licenses/by-sa/3.0/>

. In accordance with CC-BY-SA, if you distribute this document or an adaptation of it, you must provide the URL for the original version.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux ® is the registered trademark of Linus Torvalds in the United States and other countries.

Java ® is a registered trademark of Oracle and/or its affiliates.

XFS ® is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL ® is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js ® is an official trademark of Joyent. Red Hat Software Collections is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack ® Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

## 概要

本ガイドは、Red Hat Enterprise Linux での SR-IOV 実装と、Red Hat Enterprise Virtualization でのデバイス割り当てに関するハードウェアの考慮事項について説明します。

---

## 目次

1.はじめに .....	2
1.1. SR-IOV に関するハードウェア考慮事項の概要 .....	2
2. デバイス割り当てを使用するための追加のハードウェア考慮事項 .....	2
2.1. デバイス割り当て機能に関するハードウェア考慮事項の概要 .....	3
A. 改訂履歴 .....	4

## 1. はじめに

Single Root I/O Virtualization (SR-IOV) は、単一の PCI Express (PCIe) エンドポイントを複数の別個のデバイスとして使用できるようにするハードウェアリファレンスです。これは、Physical Function (PF) と Virtual Function (VF) の 2 つの PCIe 機能を導入することによって実現します。

Physical Function とは、SR-IOV 機能を含む従来の PCIe 機能で、PCIe デバイスを完全に設定/制御することができます。各 PCIe デバイスには、独立した PF を 1 つから 8 つまで使用することができます。

Virtual Function とは、軽量の PCIe 機能で、データの移動に必要なリソースと、最小限の設定リソースセットを含みます。複数の VF を各 PF 上に作成して、各 PF が異なる数の VF をサポートできるようにすることが可能です。許容される VF の合計数は、PCIe デバイスベンダーによって異なり、またデバイスによっても異なります。

PCIe の仕様は、PCIe ヘッダーのデバイス数フィールドを再換算して 9 個以上の機能を使用可能にする Alternative Routing ID Interpretation (ARI) の実装によって、より多くの VF をサポートします。この換算は、PCIe デバイスおよびそのデバイスのすぐ上位のポート (ARI をサポートするルートポートまたはスイッチ) の両方に依存します。

システムファームウェア (BIOS または UEFI) は、メモリー、I/O ポートアパーチャー、および PCIe バス番号の範囲などのリソースを割り当てます。このため、十分なリソースを割り当てるためにはファームウェアが SR-IOV をサポートおよび有効化している必要があります。

### 1.1. SR-IOV に関するハードウェア考慮事項の概要

- ファームウェアが SR-IOV をサポートしていること。
- ルートポートまたは PCIe デバイスのすぐ上位のポート (例: PCIe スイッチ) が ARI をサポートしていること。
- PCIe デバイスが SR-IOV をサポートしていること。

ハードウェアがこれらの要件を満たしているかどうかを確認するには、ベンダーの仕様とデータシートを参照してください。

`lspci -v` コマンドを使用して、システムにインストール済みの PCI デバイスの情報を表示することができます。

## 2. デバイス割り当てを使用するための追加のハードウェア考慮事項

デバイス割り当て機能により、PCIe デバイスに仮想ゲストを直接割り当てて、ゲストへの完全なアクセスとネイティブに近いパフォーマンスを提供することができます。この機能を SR-IOV と併せて実装すると、VF が仮想ゲストに直接アタッチされます。この方法で、複数の仮想ゲストを単一の PCIe デバイスの VF に直接割り当てることができます。

仮想マシンを PCIe デバイスに直接割り当てるために SR-IOV を有効にする必要はありません。また、VF 作成の用途はデバイスの割り当てだけではありません。2 つの機能は補完関係にあり、これらを併用するには、追加のハードウェア考慮事項があります。

デバイス割り当て機能には、CPU とファームウェアで I/O Memory Management Unit (IOMMU) がサポートされている必要があります。IOMMU は I/O Virtual Addresses (IOVA) と物理メモリアドレス間の変換を行

います。これにより、仮想ゲストがデバイスをゲストの物理アドレスでプログラムすることが可能となります。それらのアドレスは、IOMMU がホストの物理アドレスに変換することができます。

IOMMU グループとは、システム内の他すべてのデバイスから分離することができるデバイスのセットです。IOMMU グループは、IOMMU の粒度と、システム内の他の IOMMU グループからの分離の両方を備えたデバイスの最小のセットです。これにより IOMMU は、IOMMU グループ外にあり IOMMU の制御が及ばないデバイス間の直接メモリアクセス (DMA) を制限しつつ、IOMMU グループとのトランザクションを区別することができます。

仮想ゲストと PCIe デバイスの Virtual Function の間のトランザクションの分離は、デバイスの割り当てに必須です。PCIe およびサーバーの仕様で定義されている Access Control Service (ACS) 機能は、IOMMU グループ内で分離を維持するためのハードウェア標準です。ネイティブの ACS がない場合や、この機能が実装されていることをハードウェアベンダーが明言していない場合には、IOMMU グループのいずれかのマルチファンクションデバイスによって、IOMMU で保護されていない状態で発生する機能間のピアツーピア DMA が公開され、IOMMU グループが拡張されて、適切に分離されていない機能が含まれてしまうリスクを伴うことになります。

ネイティブの ACS サポートは、サーバーのルートポートにも推奨されます。このサポートがない場合には、それらのポートにインストールされたデバイスは、一緒にグループ化されます。ルートポートには、プロセッサベース (northbridge) とコントローラーハブベース (southbridge) の 2 種類があります。前述したように、デバイス割り当て機能を SR-IOV と併用する場合には、仮想ゲストが VF に割り当てられるので、それらのポートが ACS と ARI の両方をサポートする必要があります。

Intel の Xeon プロセッサ E5 ファミリー、Xeon プロセッサ E7 ファミリー、および High End Desktop Processors には、プロセッサベースのルートポートにおける ACS のネイティブサポートが含まれています。

Intel デバイスには、通常、コントローラーハブベースのルートポートでの ACS のネイティブサポートは含まれませんが、Red Hat Enterprise Linux 7.2 カーネルにはある程度の柔軟性があり、X99、X79、および 5 シリーズから 9 シリーズまでのチップセットのルートポートで ACS と同等の分離を有効にすることができます。

PCIe デバイスをインストールする際には、ベンダーの仕様を参照してルートポートが ACS をサポートしていることを確認した上で、プロセッサおよびコントローラーハブベースのルートポートを決定してください。

加えて、I/O トポロジー内の PCIe スイッチまたはブリッジも ACS をサポートしている必要があります。そうでない場合には、IOMMU グループが拡張される可能性があります。

## 2.1. デバイス割り当て機能に関するハードウェア考慮事項の概要

- CPU が IOMMU をサポートしていること。
- ファームウェアが IOMMU をサポートしていること。
- 使用するルートポートは、ACS または ACS と同等の機能をサポートしていること。
- PCIe デバイスとルートポート間の全 PCIe スイッチとブリッジが ACS をサポートしていること。
- PCIe デバイスが ACS をサポートしていること。

ハードウェアがこれらの要件を満たしているかどうかを確認するには、ベンダーの仕様とデータシートを参照してください。

`lspci -v` コマンドを使用して、システムにインストール済みの PCI デバイスの情報を表示することができます。

## A. 改訂履歴

**改訂 3.6-2**

**Fri 12 Feb 2016**

**Red Hat Enterprise Virtualization  
Documentation Team**

デバイスの割り当てのセクションを若干修正

**改訂 3.6-1**

**Tue 9 Feb 2016**

**Red Hat Enterprise Virtualization  
Documentation Team**

Red Hat Enterprise Virtualization 3.6 リリースの初版作成