



Red Hat Virtualization 4.4

SR-IOV 구현을 위한 하드웨어 고려 사항

Red Hat Virtualization에서 SR-IOV 구현을 위한 하드웨어 고려 사항

Red Hat Virtualization 4.4 SR-IOV 구현을 위한 하드웨어 고려 사항

Red Hat Virtualization에서 SR-IOV 구현을 위한 하드웨어 고려 사항

Enter your first name here. Enter your surname here.

Enter your organisation's name here. Enter your organisational division here.

Enter your email address here.

법적 공지

Copyright © 2022 | You need to change the HOLDER entity in the en-US/Hardware_Considerations_for_Implementing_SR-IOV.ent file |.

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, the Red Hat 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 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.

초록

이 문서에서는 Red Hat Enterprise Linux로 SR-IOV 구현과 Red Hat Virtualization을 사용한 장치 할당에 대한 하드웨어 고려 사항을 설명합니다.

차례

1장. 소개 3

1.1. SR-IOV의 하드웨어 고려 사항 3

1.1.1. 요약 3

1.2. 장치 할당에 대한 하드웨어 고려 사항 4

1.2.1. 요약 4

부록 A. 법률 알림 6

1장. 소개

다음은 Red Hat Virtualization에서 SR-IOV를 설정하고 구성하는 방법을 보여주는 일련의 주제 중 하나입니다.

- SR-IOV 구현을 위한 하드웨어 고려 사항(이 문서)
- PCI 패스스루에 대한 호스트 구성(사용자 환경에 적합한 설치 가이드를 선택).
 - 명령줄을 사용하여 Red Hat Virtualization을 셀프 호스트 엔진으로 설치
 - 로컬 데이터베이스가 있는 독립 실행형 관리자로 Red Hat Virtualization 설치
 - 원격 데이터베이스가 있는 독립 실행형 관리자로 Red Hat Virtualization 설치
- NIC에서 가상 기능 설정 편집
- vNIC 프로필에서 패스스루 활성화
- SR-IOV를 통해 가상 머신을 마이그레이션할 수 있도록 설정

1.1. SR-IOV의 하드웨어 고려 사항

SR-IOV(Single Root I/O Virtualization)는 단일 PCI Express(PCIe) 엔드포인트를 여러 개의 개별 장치로 사용할 수 있는 하드웨어 참조입니다. 이 작업은 두 PCIe 함수의 도입을 통해 수행됩니다. 물리적 기능(PF) 및 VF(가상 기능).

물리적 기능은 데이터 이동을 포함하여 PCIe 장치의 전체 구성 및 제어 기능을 갖춘 SR-IOV 기능이 포함된 기존의 PCIe 기능입니다. 각 PCIe 장치는 1~8 개의 독립 PF를 가질 수 있습니다.

가상 기능은 데이터 이동에 필요한 리소스와 최소화된 구성 리소스 세트를 포함하는 경량 PCIe 함수입니다. 각 PF에서 여러 VF를 생성할 수 있으며 각 PF에서는 다른 VF를 지원할 수 있습니다. 허용되는 총 VF 수는 PCIe 장치 공급업체에 따라 달라지며 장치마다 다릅니다.

PCIe 사양은 ARI(Alternative Routing ID Interpretation) 구현을 통해 더 많은 수의 VF를 지원합니다. PCIe 헤더의 장치 번호 필드를 다시 해석하여 8개 이상의 함수를 허용합니다. 이 변환은 루트 포트 또는 스위치와 ARI를 지원하는 장치의 PCIe 장치와 포트 즉시 업스트림에 의존합니다.

시스템 펌웨어(BIOS 또는 UEFI)는 PCIe 토폴로지를 위해 메모리, I/O 포트 집계 및 PCIe 버스 번호 범위를 포함한 리소스를 할당합니다. 따라서 충분한 리소스를 할당하려면 SR-IOV를 펌웨어에서 지원하고 활성화해야 합니다.

1.1.1. 요약

- 펌웨어(BIOS 또는 UEFI)는 SR-IOV를 지원해야 합니다. 확장 기능이 기본적으로 활성화되어 있는지 확인합니다. 그렇지 않은 경우 수동으로 활성화합니다. 이는 가상화 확장을 활성화하는 것과 유사합니다(VT-d 또는 AMD-Vi). 자세한 내용은 벤더 설명서를 참조하십시오.
- PCIe 장치(예: PCIe 스위치)의 바로 업스트림인 루트 포트 또는 포트는 ARI를 지원해야 합니다.
- PCIe 장치는 SR-IOV를 지원해야 합니다.

벤더 사양 및 데이터시트를 참조하여 하드웨어가 이러한 요구 사항을 충족하는지 확인하십시오.

lspci -v 명령을 사용하여 시스템에 이미 설치된 PCI 장치 정보를 출력할 수 있습니다.

1.2. 장치 할당에 대한 하드웨어 고려 사항

장치 할당을 사용하면 가상 게스트를 PCIe 장치에 직접 할당하여 게스트의 전체 액세스 권한을 제공하고 거의 네이티브 성능을 제공합니다. SR-IOV와 함께 구현된 가상 게스트에는 VF가 직접 할당됩니다. 이러한 방식으로 단일 PCIe 장치의 VF에 여러 가상 게스트를 직접 할당할 수 있습니다.

PCIe 장치에 가상 머신을 직접 할당하기 위해 SR-IOV를 활성화할 필요는 없으며 VF 생성을 위한 장치 할당이나 두 가지 기능이 보완적이며 함께 사용해야 하는 경우 추가 하드웨어 고려 사항이 있습니다.

장치 할당을 수행하려면 CPU 및 펌웨어에서 IOMMU(I/O Memory Management Unit) 지원이 필요합니다. IOMMU는 I/O 가상 주소(IOVA)와 실제 메모리 주소 사이를 변환합니다. 이를 통해 가상 게스트는 게스트 물리적 주소로 장치를 프로그래밍할 수 있으며, 이 주소는 IOMMU에 의해 물리적 주소를 호스팅하도록 변환됩니다.

IOMMU 그룹은 시스템의 다른 모든 장치와 격리할 수 있는 장치 집합입니다. IOMMU 그룹은 IOMMU 세분성 및 시스템 내의 다른 모든 IOMMU 그룹으로부터 격리되는 최소 장치 세트를 나타냅니다. 이를 통해 IOMMU는 IOMMU 그룹 외부의 장치와 IOMMU를 제한하는 동시에 IOMMU 그룹과의 트랜잭션을 구분할 수 있습니다.

가상 게스트와 PCIe 장치의 가상 기능 간의 트랜잭션 격리는 장치 할당의 기본 사항입니다. PCIe 및 서버 사양에 정의된 ACS(액세스 제어 서비스) 기능은 IOMMU 그룹 내에서 격리를 유지하기 위한 하드웨어 표준입니다. 기본 ACS가 없거나 하드웨어 벤더의 확인이 없으면 IOMMU 그룹 내의 모든 다기능 장치는 IOMMU 보호 외부에서 발생하는 기능 간에 피어 투 피어 DMA를 노출하고 IOMMU 그룹을 확장하여 적절한 격리가 부족한 기능을 포함할 수 있습니다.

네이티브 ACS 지원도 서버의 루트 포트에 권장됩니다. 그렇지 않으면 이러한 포트에 설치된 장치가 함께 그룹화됩니다. 두 가지 루트 포트, 프로세서 기반(northbridge) 루트 포트와 컨트롤러 허브 기반(southbridge) 루트 포트가 있습니다. 위와 같이 SR-IOV와 함께 장치 할당을 사용하고 가상 게스트가 VF에 할당되는 경우 해당 포트는 ACS 및 ARI를 모두 지원해야 합니다.

Intel의 Xeon 프로세서 E5 제품군, Xeon 프로세서 E7 제품군 및 High End Desktop Processors에는 프로세서 기반 루트 포트에서 기본 ACS 지원이 포함됩니다.

Intel Platform Controller Hub 기반(PCH) PCI Express Root Ports는 현재 ACS를 지원하지 않거나 비표준 ACS 구현을 사용하여 이러한 루트 포트를 통해 연결된 장치를 미세 분리하기 어렵습니다. 이러한 루트 포트는 대부분 ACS에 상응하는 기능을 지원합니다. Red Hat Enterprise Linux 7.3 커널에는 X79, X99, 5 시리즈에 이 ACS-equivalent 기능 활성화 및 100 시리즈 PCI Express 칩셋의 지원이 포함되어 있습니다.

루트 포트가 ACS를 지원하도록 PCIe 장치를 설치할 때 프로세서 기반 및 컨트롤러 허브 기반 루트 포트를 확인하려면 벤더 사양을 참조하십시오.

또한 I/O 토폴로지 내의 PCIe 스위치 또는 브릿지에는 ACS 지원이 필요합니다. 그렇지 않으면 IOMMU 그룹을 확장할 수 있습니다.

1.2.1. 요약

- CPU는 IOMMU(예: VT-d 또는 AMD-Vi)를 지원해야 합니다. IBM POWER8에서는 기본적으로 IOMMU를 지원합니다.
- 펌웨어가 IOMMU를 지원해야 합니다.
- 사용된 CPU 루트 포트는 ACS 또는 ACS와 동등한 기능을 지원해야 합니다.
- PCIe 장치는 ACS 또는 ACS 동등한 기능을 지원해야 합니다.

- PCIe 장치와 루트 포트 간의 모든 PCIe 스위치와 브리지는 ACS를 지원해야 합니다. 예를 들어 스위치가 ACS를 지원하지 않는 경우 해당 스위치의 모든 장치는 동일한 IOMMU 그룹을 공유하며 동일한 가상 시스템에만 할당할 수 있습니다.

벤더 사양 및 데이터시트를 참조하여 하드웨어가 이러한 요구 사항을 충족하는지 확인하십시오.

lspci -v 명령을 사용하여 시스템에 이미 설치된 PCI 장치 정보를 출력할 수 있습니다.

부록 A. 법률 알림

Copyright © 2022 Red Hat, Inc.

라이선스 라이선스 ([Creative Commons Attribution-ShareAlike 4.0 International License](#)). ([oVirt Project](#))에 대한 문서에서 파생됩니다. 이 문서 또는 문서의 수정을 배포하는 경우 원래 버전에 대한 URL을 제공해야 합니다.

수정된 버전에서는 모든 Red Hat 상표를 제거해야 합니다.

Red Hat, Red Hat Enterprise Linux, Red Hat 로고, Shadowman 로고, JBoss, OpenShift, Fedora, Infinity 로고 및 RHCE는 미국 및 기타 국가에 등록된 Red Hat, Inc.의 상표입니다.

Linux®는 미국 및 기타 국가에서 Linus Torvalds의 등록 상표입니다.

Java®는 Oracle 및/또는 그 계열사의 등록 상표입니다.

XFS®는 미국 및/또는 기타 국가에 있는 Silicon Graphics International Corp. 또는 해당 자회사의 상표입니다.

MySQL®은 미국, 유럽 연합 및 기타 국가에서 MySQL AB의 등록 상표입니다.

Node.js®는 Joyent의 공식 상표입니다. Red Hat Software Collections는 공식 Joyent Node.js 오픈 소스 또는 상용 프로젝트의 보증 대상이 아니며 공식적인 관계도 없습니다.

OpenStack® Word Mark 및 OpenStack 로고는 미국 및 기타 국가에서 OpenStack Foundation의 등록 상표/서비스표 또는 상표/서비스표이며 OpenStack Foundation의 권한에 사용됩니다. 당사는 OpenStack Foundation 또는 OpenStack 커뮤니티와 제휴 관계가 아니며 보증 또는 후원을 받지 않습니다.

기타 모든 상표는 각각 해당 소유자의 자산입니다.