

하드웨어 직접 할당 기술 기반 고품질 가상 플랫폼의 설계 및 구현

김선욱*, 김성운*, 김대원*, 오수철*, 문종배*, 조정현*, 최지혁*
 *한국전자통신연구원 SW 기반기술연구본부 서버플랫폼연구실
 e-mail : swkim99@etri.re.kr

Design and Implementation of High Quality Virtual Platform based on Direct Hardware Allocation Technology

Sun-Wook Kim*, Seong-Woon Kim*, Dae-Won Kim*, Soo-Cheol Oh*, Jong-Bae, Moon*,
 Jeong-Hyun Jo*, Ji-Hyuk Choi*
 *Dept. of Cloud Computing Research, ETRI

요 약

일반적인 가상화 플랫폼에서는 가상화 소프트웨어가 소프트웨어적으로 에뮬레이션하여 제공하는 가상 I/O 장치들을 가상 플랫폼에서 사용한다. 소프트웨어 에뮬레이션 기반 I/O 장치들을 사용하는 가상 플랫폼들은 성능이 떨어지고 고사양의 응용 프로그램을 지원할 수 없는 문제점을 가지고 있다. 본 논문에서는 해당 문제점을 극복하기 위해 PCI 기반 하드웨어 장치들을 가상 플랫폼에 직접 할당하여 고품질의 가상 플랫폼을 제공하는 기술을 설계 및 구현한다.

1. 서론

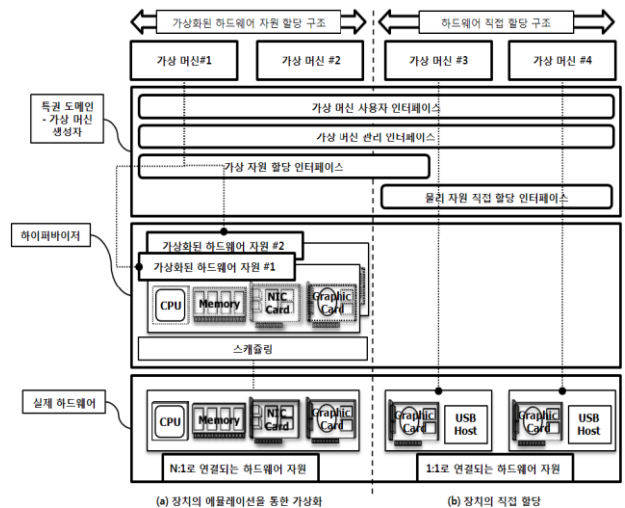
플랫폼 가상화는 효율적인 리소스 사용을 위해 물리적으로 1 대의 플랫폼을 윈도우즈나 리눅스 등 각기 다른 운영 체제 기반 가상 머신에서 공유하는 것을 말한다. 하지만 여기에서 플랫폼은 단순히 프로세서만을 의미하는 것이 아니라 스토리지, 네트워크 및 기타 하드웨어 리소스를 비롯하여 플랫폼을 구성하는 다른 중요 요소까지도 포함된 의미이다[1].

전 가상화 기반의 윈도우즈 운영체제를 지원하는 가상 머신은 통상 HVM(Hardware Virtual Machine)[7]이라 할 수 있다. HVM 환경은 기본적으로 에뮬레이션된 하드웨어 장치를 이용하여 N 개의 가상 머신이 유일한 하드웨어 장치를 공유하여 사용한다. 즉, 가상 머신에서 하드웨어 장치에 접근하려는 I/O 들은 소프트웨어로 모델링된 가상 하드웨어 자원으로 저장되고, 특권 도메인에 설치된 드라이버 모듈(실제 장치에 접근할 수 있는 S/W 모듈)을 통해 해당 I/O 들이 처리된다. 이 I/O 들을 처리하기 위한 스케줄링의 주체는 하이퍼바이저가 되고, 가상 머신들의 I/O 데이터가 집중되게 되면 모든 가상 머신의 I/O 반응 속도 저하를 초래할 수 있다.

본 논문에서는 윈도우 운영체제 기반의 가상 플랫폼에 고사양의 응용 프로그램을 구동시키기 위한 PCI

기반 하드웨어 직접 할당 기술 기반 고품질 가상 플랫폼을 설계 및 구현한다.

하드웨어 직접 할당 기반 고품질의 가상 플랫폼은 장치 에뮬레이션을 통한 I/O 오버헤드의 증가 및 이로 인한 성능 저하를 방지하기 위해 특정 가상 플랫폼에 VGA 나 USB 호스트와 같은 I/O 하드웨어 들을 선점할 수 있는 기능을 제공함으로써 가상 플랫폼에서 고사양의 응용 프로그램 구동이 가능하도록 한다.



(그림 1) 플랫폼 가상화 지원 시스템 구조

2. 하드웨어 직접 할당 기반 고품질 가상 플랫폼의 설계

2.1 하드웨어 직접 할당 기술

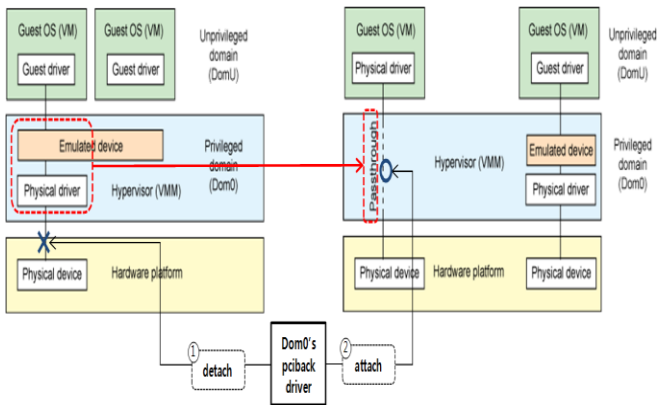
가상 플랫폼내에서 고사양 응용 프로그램을 제공하기 위해서는 각각의 사용자가 사용할 가상 플랫폼, 즉 가상 머신에게 물리적 자원을 직접 할당하는 기능을 제공하여야 한다. 이를 지원하기 위해서는 해당 물리적 하드웨어 장치의 가상 머신 접근 기능 및 해당 자원의 가상 머신에 의한 선점 기능을 제공해야 한다.

- 선점된 물리적 자원의 가상 머신 접근 기능

플랫폼 가상화 환경에서 메모리 가상화를 통해서 각 가상 머신은 자신의 독립적인 가상 주소 공간을 사용하게 되는데 가상 머신에서 물리적인 자원과 가상 머신의 1:1 연결을 위해서는 이 하드웨어에서 가상 머신의 가상 주소를 접근할 수 있어야 한다. 따라서, 인텔의 VT-D[2] 또는 AMD의 IOMMU 등으로 하드웨어에게 가상 머신에 접근할 수 있는 방법을 제공한다. 그러나, 메인 보드의 바이오스와 가상화 커널(즉, 하이퍼바이저)에서 해당 기능을 완벽히 지원하지 않는다. 따라서, 하이퍼바이저를 수정하여 VT-D 및 IOMMU 기능을 지원하도록 제공한다.

- 고사양 응용 프로그램 지원을 위한 물리적 자원 선점 기능

(그림 2)는 고사양 응용 프로그램을 위한 PCI Passthrough[3] 설정 구조를 보여준다. (그림 2)의 pci_back 드라이버를 통해서 특권 도메인(Dom0)에서 선점하고 있는 물리적 하드웨어 자원의 연결 패스를 제거할 수 있다. 또한, 하이퍼바이저는 pci_back 드라이버를 이용하여 특정 가상 머신에게 물리적 PCI 장치를 선점할 수 있도록 제공한다.



(그림 2) 고사양 응용 프로그램 지원을 위한 PCI Pass-through 설정 구조

- 펌웨어 연동 하드웨어 장치의 직접 할당

네트워크 카드와 USB 호스트 같은 일반적인 PCI 장치와 달리 그래픽 카드는 더 복잡한 구조를 가지고 있기 때문에 일반적인 PCI 장치(네트워크 카드, USB 등)의 Pass-through에 추가적인 기술이 요구된다[4,5].

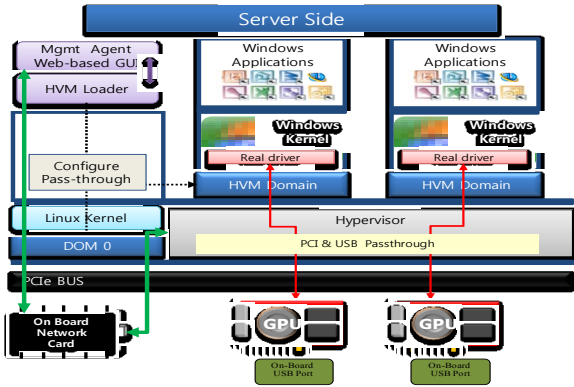
VGA 카드는 부팅 시 사용자 인터페이스를 제공하기 위해서, 부팅 전 VGA 카드의 롬에 저장된 펌웨어 프로그램을 확장 VGA 바이오스 영역에 복사하여 이를 통해 VGA 장치를 초기화 한다. 또한, 윈도우 그래픽 환경에서 3D와 같은 고품질 그래픽을 표현하기 위해서 부팅 이후 해당 VGA 장치에 종속적인 드라이버 프로그램이 추가적으로 설치되어야 한다. 기본적으로, 하이퍼바이저의 VGA 에뮬레이션은 가상화된 VGA 카드에 상응하는 VGA 드라이버 펌웨어를 가상 머신의 VGA 바이오스 영역에 복사하고, 이 둘 사이의 인터페이스를 통해서 N 대 가상 머신의 그래픽을 표현한다. 에뮬레이션된 VGA 카드는 실제 VGA 카드 종속적인 드라이버를 설치할 수 없기 때문에 실제 VGA 카드 종속적인 DirectX와 같은 고급 기능을 사용할 수 없다. 다음은 VGA 카드의 Pass-through 기반 직접 할당 기술 방법을 나타낸다.

- 가상화 서버 플랫폼에 장착된 실제 VGA 카드에 상응하는 펌웨어를 가상 머신의 VGA 바이오스 영역에 복사해준다.
- 가상 데스크탑 생성 블록에는 근거리 DaaS의 VGA Pass-Through 설정 시, 가상 머신 생성자에서 기본적으로 제공하는 VGA 가상화 기능을 비활성화 시킨다.
- VGA 장치 제조사 별로 혹은 VGA 장치가 장착되는 PCI/PCIe 슬롯에 따라서 VGA 장치 BAR(Base Address Register)가 연결되는 메모리 주소 영역이 다르다. 따라서, 가상 데스크탑 생성 블록에서 가상 머신의 가상 주소 공간과 실제 VGA 장치 BAR가 연결되는 메모리 주소를 재 연결 해주는 기능을 제공한다.

2.2 하드웨어 직접 할당 기반 고품질 가상 플랫폼의 구조

하드웨어 직접 할당 기반 고품질 가상 플랫폼은 3D 게임 및 고품질의 동영상 재생과 같은 높은 수준의 하드웨어 사양이 요구되는 애플리케이션 서비스를 위한 구조이다. 즉, 고품질 가상 플랫폼의 구조에서는 일반 가상 플랫폼에서 복수의 가상 머신이 특권 도메인의 중재를 통해 N:1로 하드웨어 자원과 연결되는 패스를 제거하고, 가상 머신과 하드웨어 자원이 1:1로 연결된다(Pass-Through). 따라서, 고품질 가상 플랫폼용 가상 머신이 생성되는 시점에서 특권 도메인[6]

은 자신이 선점하던 하드웨어 자원을 반납하고, 이를 생성된 가상 머신과 연결시킨다. 가상 머신과 직접적으로 연결된 하드웨어 자원은 PCI/PCIe 그래픽 카드 및 USB 키보드, USB 마우스 등이 있다.



(그림 3) 하드웨어 직접 할당 기반 고품질 가상 플랫폼 구조

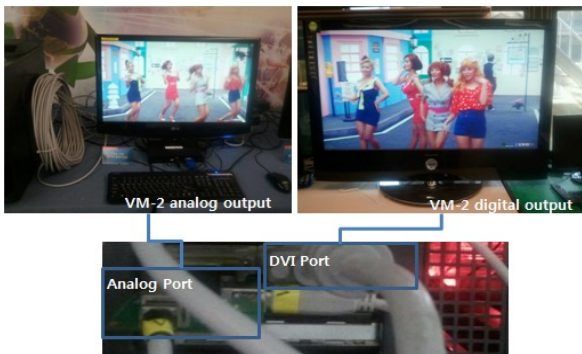
3. 구현

하드웨어 직접 할당 기반 고품질 가상 플랫폼 시스템은 Xen 하이퍼바이저를 기반으로 구현되었다. 해당 플랫폼은 VT-d 기능을 제공하는 프로세서 및 메인보드를 사용하였으며 하나의 플랫폼에 두 개의 VGA 카드를 장착하여 각 가상 머신에 직접 할당 하였다.

또한 사용자 인터페이스 제공을 위한 USB Host Controller 를 각 가상 머신에 직접 할당 하였다. 각 가상 머신은 윈도우 7 을 기반으로 하며 DirectX 를 기본적으로 사용하는 테스트 프로그램을 설치하였다.



(그림 4) 고품질 가상 플랫폼의 3D 게임 지원



(그림 5) 고품질 가상 플랫폼의 듀얼 모니터 기능 지원

(그림 4)는 구현된 고품질 가상 플랫폼에서 3D 기반의 게임을 동시에 구동하는 모습을 나타내고 (그림 5)는 듀얼 모니터를 지원하는 가상 플랫폼에서 Full-HD 동영상 재생을 나타낸다.

일반 데스크탑과 2 대의 가상 머신에서 동시 수행된 3D 성능 측정 벤치마크 결과는 아래 표와 같다. 해당 수치는 각각 10 번씩 측정된 결과의 합을 바탕으로 작성되었으며 일반 데스크탑 대비 165.58%의 성능을 보이는 것을 알 수 있다. 이는 VGA 카드를 각각의 가상 플랫폼에게 할당함으로써 한 대의 물리적 플랫폼에서 두 개의 가상 플랫폼 제공을 통해 고사양의 어플리케이션 구동을 가능하게 한다.

<표 1> 2D/3D 기반 VGA 성능 측정 결과

	Native(point)	2VM(point)	Relative Ratio(%)	Test Result
3DMark Score	16541	27388	165.58	Pass

4. 결론

하드웨어 직접 할당 기반 고품질 가상 플랫폼은 장치 에뮬레이션 기반 가상 플랫폼에서 발생하는 I/O 오버헤드의 증가 및 이로 인한 성능 저하를 방지하기 위해 특정 가상 플랫폼에 VGA 나 USB 호스트와 같은 I/O 하드웨어 들을 선점 및 할당 할 수 있는 기능을 제공함으로써 가상 플랫폼에서 고사양의 응용 프로그램 구동이 가능하게 해준다. 해당 기술은 현재 입출력 가상화 기반 네트워크 하드웨어의 발전으로 다수의 가상 플랫폼에 동일한 성능의 네트워크 성능을 제공하고 있다.

참고문헌

- [1] Bencang Liu et al., "Research on Hardware I/O Passthrough in Computer Virtualization," Proc.of ISCSCT 2010, Aug. 2010, pp. 353~356.
- [2] Xen wiki, VTdHowTo, "http://wiki.xensource.com/xenwiki/VTdHowTo"
- [3] LIU, J et al, "High Performance vmm-bypass i/o in virtual machines," Proc. of USENIX Annual Technical Conference, May. 2006.
- [4] Beng Heng Ng et al, Direct Access to Graphics Card Leveraging VT-d Technical Report, http://www.eecs.umich.edu/~bengheng/pubs/vgapt_techreport.pdf"
- [5] Y. Schaeffer, Xen VGA passthrough, http://staff.science.uva.nl/~delaat/sne-2008-2009/p22/report.pdf"
- [6] Xen wiki, XenDom0Kernels, http://wiki.xensource.com/xenwiki/XenDom0Kernels
- [7] BARHAM, P. et al, "Xen and the art of virtualization," Proc.of SOSP, Oct.2003.