

# PCIe 확장 보드용 통신 매체에 대한 응용 레벨 성능 평가

차광호, 구경모, 정현미  
 한국과학기술정보연구원 슈퍼컴퓨터기술개발센터  
 e-mail: khocha@kisti.re.kr

## Application-level Performance Evaluation of Transmission Media for PCIe Expansion

Kwangho CHA, Kyungmo Koo, Hyun Mi Jung  
 Div. of Supercomputing, Center for Developement of Supercomputing System  
 Korea Institute of Science and Technology Information

### 요약

PCI Express(PCIe) 버스는 시스템 내·외부의 다양한 디바이스들을 연결하는데 여전히 중요한 역할을 하고 있다. 이 PCIe 버스를 확장하기 위해서 PCIe 스위치가 사용되며 호스트(서버)와 외부 디바이스를 연결할 때 다양한 케이블 미디어가 사용된다. 본 연구에서는 광 타입과 구리선 타입의 2종류 케이블이 장착 가능한 자체 제작된 PCIe 어댑터 카드를 사용하여 응용 프로그램 레벨에서 성능을 확인하였다. 서로 다른 2종류의 디바이스들을 대상으로 실험한 결과, 성능상의 큰 차이는 발견되지 않아서 케이블의 종류와는 무관하게 상호 운영에 크게 문제가 없을 것으로 예상된다.

### 1. 서론

시스템 내·외부의 다양한 디바이스들을 연결하는데 PCI Express(PCIe) 버스는 여전히 중요한 기술이며 이 버스의 적용 범위를 넓히는데 PCIe 스위치가 사용된다. 다시 말해서, 호스트(서버) 외부의 디바이스를 연결하거나 호스트들을 연결하는데 PCIe 버스가 사용되고 있다[1,2]. 특히 GPU나 외장 SSD와 같은 외부 장치들의 활용도가 높아지면서 PCIe 스위치를 활용한 외부 확장 보드들의 관심도 높아지는 상황이다[3]. 이러한 외부 확장 보드를 서버와 연결할 때 일반적으로 HIF(Host Interface)카드라고 하는 일종의 어댑터 카드를 사용하게 되며 이는 다시 PCIe 신호를 전송할 수 있는 케이블로 연결된다.

본 논문에서는 자체 제작된 PCIe 어댑터 카드[4]에 광 케이블과 구리선 케이블을 교대로 장착하면서 디바이스들의 연결상 특이점을 살펴보았다. 자체 제작한 PEX9797 기반의 확장 보드에 GPU와 SSD를 장착하여 응용 프로그램 레벨에서의 성능을 측정하였다.

### 2. 온보드형 광모듈 기반 어댑터 카드

실험에 사용된 어댑터 카드는 Broadcom社의 PEX 8733을 PCIe 스위치로 사용하며 소형화된 온보드형 광모듈을 활용하여 PCIe 버스 신호를 전송하는 형태로 제작되었다. 어댑터 카드의 구성을 변경하면 PCIe 버스 확장 이외의 호스트간 통신에도 사용 가능하다[4].

사용된 광 모듈은 착탈 가능한 형태여서 광 모듈부 대신 광모듈 제조사에서 제공하는 구리선 기반의 케이블 모듈을

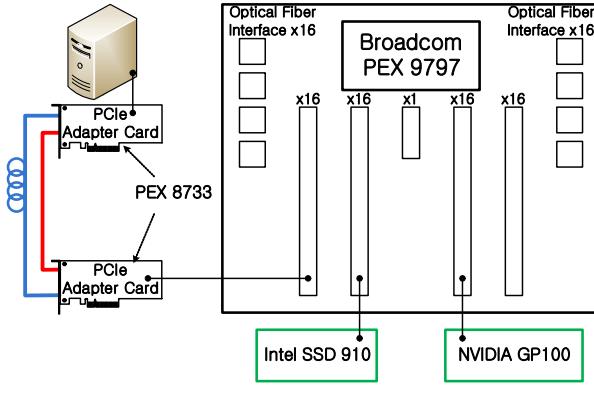
<표 1> 전송 모듈[Samtec社]별 주요 스펙

모듈명	케이블 종류	모듈당 채널수	유효전송 거리	인터페이스
PCUO	광	x4(x8)	100M	UEC5, UCC8, OPA
PCUE	구리선	x4	10M	UEC5, UCC8

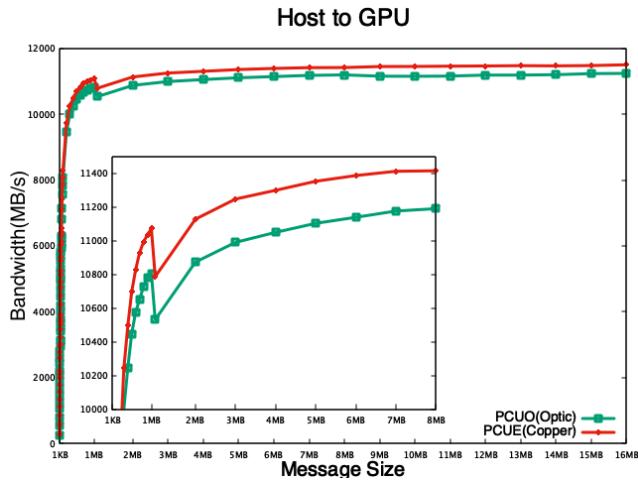
바로 장착할 수 있는 구조이다. 표 1은 실험에 사용된 두 가지 전송 모듈들의 특징을 요약한 것이다[5,6]. 두 전송 모듈 모두 동일한 형태의 물리적인 인터페이스를 사용하고 있으나 유효 전송 거리에서 차이를 보이고 있다. 공통적으로 모듈 당 PCIe Gen 3 규격의 4채널 지원이 가능하다. 자체 제작된 PCIe 어댑터 카드는 4개의 모듈을 장착하여 16배속 통신이 가능하도록 하였고 본 실험에서도 물리 계층의 연결은 전송 매체별로 16배속을 확보한 뒤 성능을 측정하였다.

### 3. 실험 환경

PCIe 버스 확장시 전송 케이블의 특성 확인을 위해 그림1과 같은 실험환경을 구축하였다. 앞서 기술된 PCIe 어댑터 카드를 서버(다운스트림 포트)와 확장 보드(업스트림 포트)에 각각 장착하였다. 이때 확장 보드에 내포된 광 모듈부를 업스트림 포트로 활용하여 서버와 연결할 수도 있으나 확장 보드로부터 발생할 수 있는 영향을 배제하기 위하여 어댑터 카드 2개를 활용하여 연결하는 방식을 취하였다.



(그림 1) 실험 환경



(그림 2) CPU와 GPU간 대역폭

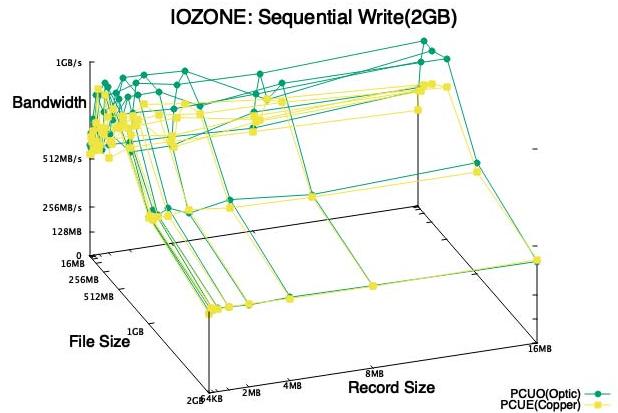
확장 보드는 Broadcom社의 PEX 9797을 사용하여 자체 제작하였다. Gen 3 PCIe 96레이인의 지원이 가능해서 4개의 16레이인 표준 슬롯과 광 모듈에 32레이인을 할당한 구조이다.

다운스트림 포트에 GPU와 SSD를 설치하고 서버에서 각 장치를 접근할 때의 대역폭을 측정하였다. 이때 구리선 타입은 1미터, 광 타입은 10미터 길이의 케이블을 각각 사용하였다.

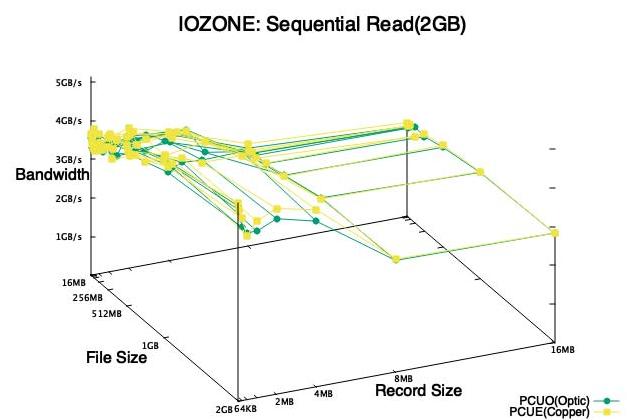
#### 4. 성능 측정 결과

확장 보드에 장착된 GPU를 사용할 때의 대역폭을 살펴보았으며 CUDA Tool kit의 대역폭 측정 프로그램을 활용하였다. 정확한 대역폭을 측정하기에는 한계가 있다고 할 수 있으나 두 경우에 대한 상대적인 성능 비교는 가능하다고 판단하였다.

그림 2는 하나의 GPU로 측정한 대역폭으로 16배속의 PCIe 대역폭을 모두 채우기에는 한계가 있었으나 확장 보드에 복수의 GPU를 장착하는 경우에는 16배속 대역폭을 모두 채울 것으로 예상하고 있다. 그러나 이런 경우에는 두 종류의 통신 매체의 차이점을 구분하기 어려운 면이 있기 때문에 한 개의 GPU를 사용하여 대역폭을 측정하였다. 실험 결과는 구리선 타입의 케이블을 사용했을 때 광



(그림 3) IOZONE 쓰기 대역폭



(그림 4) IOZONE 읽기 대역폭

케이블 보다 약 2% 정도 높은 대역폭을 보여주었다.

두 번째 실험에서는 확장 보드에 PCIe 인터페이스 기반의 SSD를 설치하고 ext3 파일 시스템을 구성하여 IOZONE을 통한 IO 성능을 측정하였다. 2GB 파일을 쓰고 읽는 성능을 그림 3과 4에서 확인할 수 있다. 쓰기의 경우, 최대 약 1GB/s의 성능을 보였으며 읽기에서는 최대 5GB/s의 성능을 확인할 수 있었으며 쓰기의 일부 구간에서 광 타입의 케이블이 구리선 타입 케이블 보다 상회하는 성능을 보여주었다. 그러나 그 외의 경우에는 두 종류 케이블의 우열을 가리기 어려운 상황이었다.

#### 5. 결론 및 향후 계획

본 연구에서는 온 보드형 광모듈 기반의 PCIe 어댑터 카드와 확장 보드를 사용하여 PCIe 버스를 확장한 후, 전송 미디어가 변경되는 경우의 영향을 살펴보았다. 즉, 호스트 서버와 확장 보드의 업스트림 포트를 연결하는 미디어를 광 타입과 구리선 타입 두 종류를 사용하여 성능을 측정하였다. 단순히 통신 대역폭을 측정하는 경우에는 구리선 타입의 케이블이 약간 좋은 성능을 보여 주었으나 전반적

으로는 두 경우의 차이를 명확히 언급하기 어려웠다. 결국 응용프로그램을 사용한 성능 측정에서는 두 경우의 급격한 성능 차이가 나타나지 않았기 때문에 구성하고자 하는 시스템의 성격, 즉, 가격과 연결 거리 등을 감안하여 선택하는 것이 옳을 것으로 판단된다. 향후 확장 보드의 특성과 결합되는 성능 상 특성을 세밀하게 분석할 계획이다.

### Acknowledgement

본 연구는 2019년도 한국과학기술정보연구원(KISTI) 주요 사업 과제(K-19-L02-C06-S01)로 수행되었습니다.

### 참고문헌

- [1] Cheol Shim, Kwangho Cha, Min Choi, "Design and implementation of initial OpenSHMEM on PCIe NTB based cloud computing," Cluster Computing, vol. 22, no. 1 supplement, pp. s1815~s1826, Jan. 2019.
- [2] Seung-Ho Lim, Ki-Woong Park, and Kwangho Cha, "Developing an OpenSHMEM model over a Switchless PCIe Non-Transparent Bridge Interface," IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), pp. 593~602, May. 2019.
- [3] One stop system, <https://www.onestopsystems.com/expansion-backplanes>
- [4] Kyungmo Koo, Junglok Yu, Sangwan Kim, Min Choi, and Kwangho Cha, "Implementation of Multipurpose PCI Express Adapter Cards with On-Board Optical Module," The Journal of Information Processing Systems, vol. 14, no. 1, pp. 270~279, Feb. 2018.
- [5] PCIe®-OVER-FIREFLY™ COPPER FLYOVERTM,  
<https://www.samtec.com>
- [6] PCIe® - OVER - FIBER OPTICAL FLYOVERTM,  
<https://www.samtec.com>