

인피니밴드 인메모리 스토리지 구현 및 성능평가

성창경^o
 단국대학교 컴퓨터학과 컴퓨터공학^o
 e-mail: cgseong@dankook.ac.kr^o

Implementation of InfiniBand In-memory Storage and Performance Evaluation

Chang-Gyeong Seong^o
 Dept. of Computer Engineering, Dankook University^o

● 요약 ●

본 논문은 퍼스널 컴퓨팅 환경의 성능 향상을 위한 인피니밴드 네트워크 기반 인메모리 스토리지 시스템의 구조를 제안한다. 성능평가를 위해 100Gbit/s를 지원하는 MCX455A-ECAT 한 쌍을 MCP1600-E02A 케이블로 직결한 x86-64 architecture의 인피니밴드 네트워크를 구성하고 Target 시스템에 iSCSI Extensions for RDMA(iSER)을 적용한 RAM disk를 생성하였다. CentOS virt-manager에서 생성한 Initiator 시스템의 Windows 가상 머신에는 Target 시스템의 RAM disk를 VirtIO 방식으로 연결한다. 이 구조는 시스템 종료 후 초기화되는 종래 RAM disk의 일반적 특성을 개선한다. 마지막으로 스토리지 성능평가를 통해 향후 출시될 PCI Express 4.0 이상의 시스템과 퍼스널 컴퓨팅 스토리지 성능 향상 측면에서 해당 구조의 적합성을 보인다.

키워드: 인피니밴드(InfiniBand), 인메모리 스토리지(In-memory Storage), virt-manager, VirtIO, RDMA

I. Introduction

Solid-state drive(SSD)의 등장[1]과 성능 향상[2]으로 SSD가 Hard disk drive(HDD)를 빠르게 대체하고 있다. 이러한 변화에 따라 SSD를 상회하는 성능의 저장장치가 요구되고 있다.

종래의 RAM disk는 일반적으로 SSD보다 월등한 성능을 갖고 있지만 운영체제에 종속된 구조이며 시스템 구동 중의 모든 프로세스는 RAM disk가 사용하는 Bus를 공유하게 되어 필연적으로 병목현상이 발생한다. 또한 시스템 종료 후에 저장된 데이터가 모두 초기화되는 단점이 있었다.

본 논문에서는 인피니밴드 네트워크 기반 인메모리 스토리지 구조를 제안하고 실제 시스템 구현과 성능평가를 통해 해당 구조의 적합성을 확인한다.

II. Preliminaries

1. Mellanox ConnectX[®]-4 VPI Card

스토리지 네트워크 구현에 사용된 100Gbit/s 인피니밴드, 이더넷 겸용 Adapter Card이다. Mellanox Technologies 사의 MCX455A-ECAT 장비 2점과 MCP1600-E02A 케이블 1점이 구현에 사용되었다. 해당 케이블은 IB EDR, 100Gbit/s를 지원한다.

2. iSCSI Extensions for RDMA (iSER)

RDMA는 CPU를 사용하지 않고 직접 메모리에 읽고 쓸 수 있는 기술이다. iSER[3]은 RDMA를 활용한 iSCSI 구현으로 전통적인 iSCSI 보다 향상된 성능을 보인다. 인메모리, SSD 스토리지 모두 iSER를 적용하여 성능평가를 실시했다.

III. The Proposed Scheme

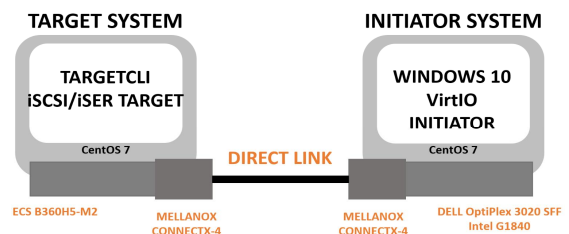


Fig. 1. System Architecture

Table 1. Target System Environment

Item	Value
프로세서	Intel Celeron G4900
설치된 메모리(RAM)	8.0GB
시스템 모델	ECS B360H5-M2

퍼스널 컴퓨팅 환경의 성능 향상을 위한 인피니밴드 네트워크 기반 인메모리 스토리지 시스템 구조의 구현 및 성능평가는 크게 여섯 가지 순서로 구성된다.

- MCX455A-ECAT 모델이 설치된 Target 시스템과 Initiator 시스템을 MCP1600-E02A 케이블로 직결한 후 양쪽에 CentOS 7 운영체제를 설치한다.

- Target 시스템의 RAM 4GB를 targetcli을 통해 iSCSI Target RAM disk로 생성한다.

- Initiator 시스템의 virt-manager를 사용하여 생성한 가상 머신에 Windows 10을 설치한다.

- iSCSI Target을 Initiator 시스템의 CentOS 운영체제 환경에서 연결한다.

- Initiator 시스템의 Windows 10 가상 머신에서 앞서 iSCSI으로 연결된 RAM disk를 VirtIO 방식으로 연결한다.

- Windows 10 가상 머신에서 DiskSpd Utility를 사용하여 인메모리 스토리지의 성능평가를 시행하고, 부가적으로 RAM disk를 Intel SSDSC2BW120H6로 교체한 동일 구조에서 성능평가를 시행한다 [4][5].

시행되었기 때문에 완전한 PCIe 3.0 시스템에서는 설치된 RAM의 대역폭에 근접한 성능이 예상된다. 따라서 향후 출시될 PCI Express 4.0 규격 시스템에서 해당 구조의 실험을 통해 스토리지 성능 개선 연구를 진행할 예정이다.

REFERENCES

- [1] 1991: Solid State Drive module demonstrated, <https://www.computerhistory.org/storageengine/solid-state-drive-module-demonstrated/>
- [2] Specifications - NVM Express, <https://nvmexpress.org/resources/specifications/>
- [3] Mike Ko, "iSCSI Extensions for RDMA Specification (Version 1.0)," RDMA Consortium, pp.1-14, Jul. 2003.
- [4] Hui-Seong Heo, Kwang-Soo Lee, Mehdi Pirahandeh, Deok-Hwan Kim, "Design of OpenStack Cloud Storage Systems - Applying Infiniband Storage Network and Storage Virtualization Performance Evaluation," KIISE Transactions on Computing Practices, Vol. 21, No. 7, pp. 470-475, Jul. 2015.
- [5] Hui-Seong Heo, Mehdi Pirahandeh, Kwang-Soo Lee, Deok-Hwan Kim, "All Flash Array Storage Virtualisation using SCST," KIISE Transactions on Computing Practices, Vol. 20, No. 10, pp. 525-533, Oct. 2014.

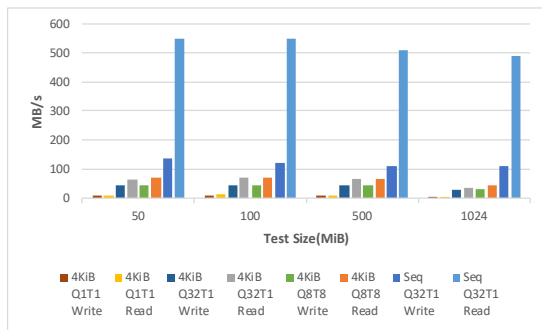


Fig. 2. I/O Performance of iSER SSD Storage

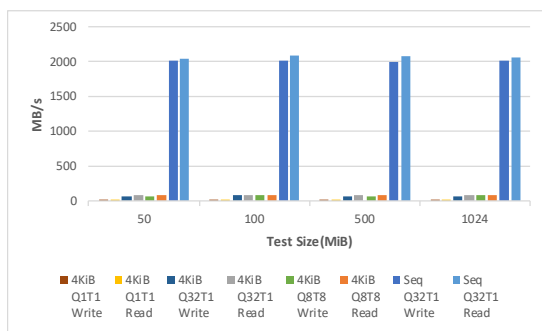


Fig. 3. I/O Performance of iSER In-memory Storage

IV. Conclusions

본 논문에서는 퍼스널 컴퓨팅 시스템에 적용 가능한 인피니밴드 네트워크 기반 인메모리 스토리지 시스템 구조를 제안하였고, 성능평가를 통해 해당 구조의 타당성을 보였다. 저사양 CPU, PCIe 2.0 수준 대역폭의 Bus로 구성된 제한적 성능의 시스템에서 성능평가