

Infrastructure-as-a-Service의 Disaggregated System 내 Storage Node SSD Trace 특성에 따른 최적화

*성명호, 정의영

연세대학교 전기전자공학부 소속

e-mail : smh@yonsei.ac.kr, eychung@yonsei.ac.kr

Optimization of Storage Nodes in Disaggregated Systems within Infrastructure-as-a-Service Based on SSD Trace Characteristics

*MyeongHo Seong and Eui-Young Chung
School of Electrical and Electronic Engineering
Yonsei University

Abstract

In this paper, we explore SSD I/O trace characteristics within a disaggregated system, comparing them against traditional aggregated systems. For disaggregated systems, we introduce the "Gradual Lazy Garbage Collection", designed to optimize block erase distribution and extend SSD lifespan by adjusting the garbage collection threshold based on system status. Our findings demonstrate that this method effectively balances wear distribution and enhances the longevity of storage devices in IaaS cloud environments. Further research will focus on expanding target traces to address wear leveling challenges and develop advanced garbage collection strategies.

I. 서론

기술 발전과 고속 네트워크의 대중화로 클라우드 컴퓨팅이 상용화되고 있다. 클라우드 컴퓨팅은 위치 제약 없이 뛰어난 접근성, 확장성, 그리고 유연성을 제공하며 이를 통해 경쟁력을 확보하였다. 특히 IaaS

(Infrastructure as a Service)는 서버, 스토리지, 네트워크 자원을 포함한 컴퓨팅 하드웨어 리소스를 가상화하여 제공하는 서비스로 주목받고 있다. 대표적으로 Amazon의 AWS, NHN의 Cloud Instance Service가 있다.

IaaS는 사용자에게 PaaS (Platform as a Service)나 SaaS (Software as a Service)와 비교하여 더 광범위한 시스템 제어 권한과 자유도를 제공한다. 이를 통해 사용자는 할당받은 자원을 이용하여 자신만의 가상 환경에서 다양한 애플리케이션과 서비스를 구축할 수 있다. 이와 같은 특성 덕분에 IaaS 환경에서 자원 할당과 서비스 관리에 관한 많은 연구가 이루어지고 있다.[1][2]

IaaS의 disaggregated 환경에서 compute node는 사용자의 작업 부하를 처리하며, 이로 인한 data I/O는 storage node의 SSD에 기록된다. 이 구조는 compute 자원과 storage 자원이 통합된 전통적인 aggregated 환경과 다른 I/O trace를 생성한다. 본 논문에서는 disaggregated system에서 나타나는 SSD I/O trace의 특성을 aggregated system에서의 결과와 비교 분석한다. 이러한 분석을 통해 storage node의 SSD에서 발생하는 trace가 어떻게 처리되는지 파악하고, 이러한 시스템에서 SSD의 수명에 미치는 영향을 논의한다. 또한 시스템에서 SSD의 수명을 관리하기

위한 효과적인 garbage collection 방법을 제안한다.

II. 실험 방법

2.1 IaaS 환경의 I/O특성 분석을 위한 데이터

그림 1과 같이 IaaS 환경을 모사한 compute node와 storage node로 구성된 시스템에서 사용자가 작업 공간을 반복적으로 생성 및 소멸하는 시나리오를 수행하였다. 이 과정에서 발생하는 storage node의 SSD 접근 trace 데이터를 수집하였고, 이 데이터를 분석하여 IaaS환경에서의 주요 I/O 특성을 확인하였다. 이를 통해 aggregated system과 disaggregated system에서의 데이터를 비교 분석하였다.

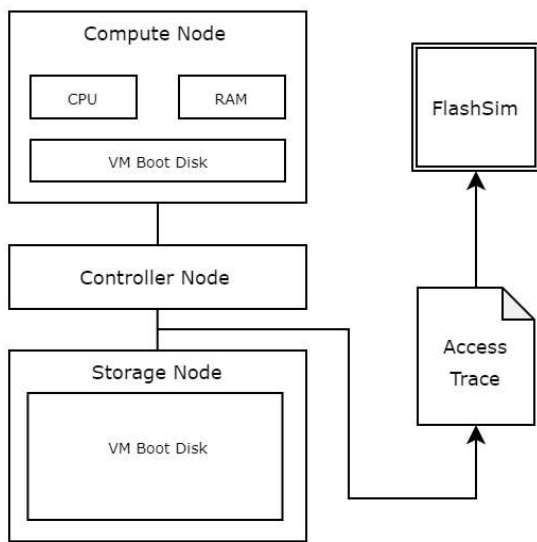


그림 1. Disaggregated System Trace 추출 및 분석

2.1 I/O 특성 분석과 SSD에 미치는 영향 및 적절한 Garbage Collection 방법 제안

본 연구에서는 FlashSim[3] 기반의 trace-driven SSD simulator를 사용하여 64GB SSD에서의 데이터 처리 방식을 분석하였다. 이 SSD는 DFTL(Demand Based Flash Translation Layer) 기법[4]을 기반으로 관리된다. 실험 과정에서 전체 block 대비 write가 불가능한 block (유효 혹은 무효 상태)의 비율을 지속적으로 모니터링하여 이 비율이 특정 임계값을 초과할 경우 지정된 수의 erase 대상 block을 garbage collection을 통해 처리하도록 설계하였다. Event 진행에 따라 임계값과 erase 대상 block수를 점진적으로 조정하는 'Gradual Lazy Garbage Collection'을 제시하여 효과를 확인해보았다.

III. 실험 결과

그림 2는 전통적인 aggregated 환경에서 수집된 logical write address 분포이며, 그림 3에서는 수집된 trace 데이터를 통해 그림 2에 비해 특정 logical address에 대한 random write 작업이 자주 발생함을 확인할 수 있다.

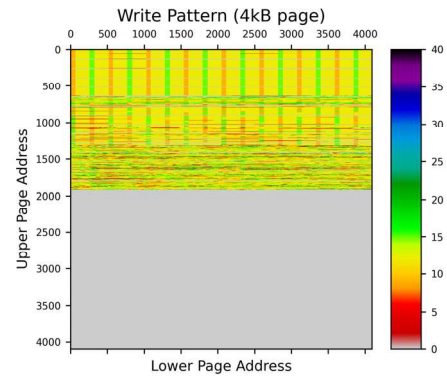


그림 2. Aggregated System에서의 Logical Write Address 분포

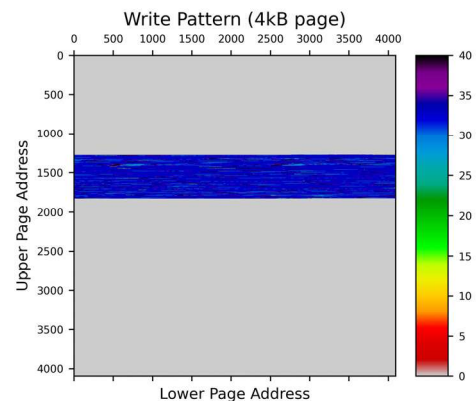


그림 3. Disaggregated System에서의 Logical Write Address 분포

이는 IaaS 환경에서 사용하는 QCOW2 이미지의 특성 [5]과 관련이 있다. 기본 OS image를 제외한 변경 사항만이 storage node에 기록되기 때문이며, 특히 이미지의 헤더 부분에 write 작업이 반복적으로 일어나는 것의 영향이 크다고 할 수 있다.

그림 4는 FlashSim을 통해 구성된 DFTL과 'Gradual Lazy Garbage Collection'을 사용하는 64GB SSD에서 입력된 trace 데이터로부터 계산된 물리적인 block별 erase분포를 보여준다. 해당 결과는 실제 SSD에서 데이터가 어떻게 관리되는지에 대한 중요한 지표이다.

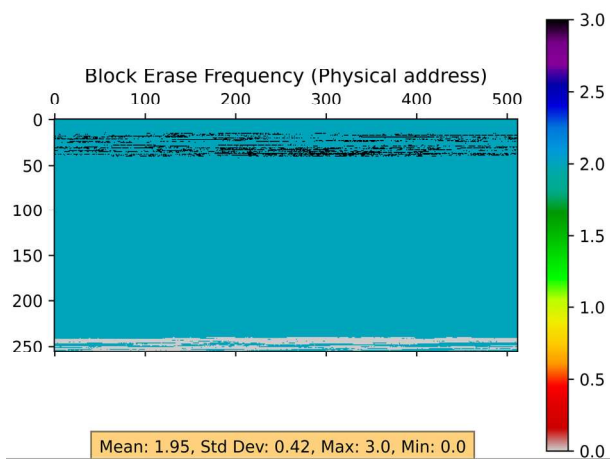


그림 4. Block Erase Count 분포

그림 4에 따르면, 전체 block의 2.7%에서 세 번의 erase 작업이 일어났으며, 대다수인 93.6%의 block에서는 2번의 erase 작업이 이루어졌고, 3.7%의 block에서는 erase 작업이 일어나지 않았다. Erase가 일어나지 않은 block은 주로 사용 빈도가 낮은 cold data가 저장되어 있어 erase count가 증가하지 않은 valid block일 가능성이 높다. 나머지 block들에서는 고르게 write 작업이 진행되었으며, 이는 순차적으로 오래된 invalid block을 erase함으로써 block의 마모 횟수가 고르게 관리되었다는 것을 시사한다. 이에 따라, 추후 작업이 진행되더라도 적은 비율의 cold block에 대해서만 별도의 static wear leveling만을 고려하는 것만으로 충분하다는 결론을 내릴 수 있다. 이러한 접근 방식은 wear leveling의 오버헤드와 write amplification을 현저히 줄이는 효과를 기대할 수 있다.

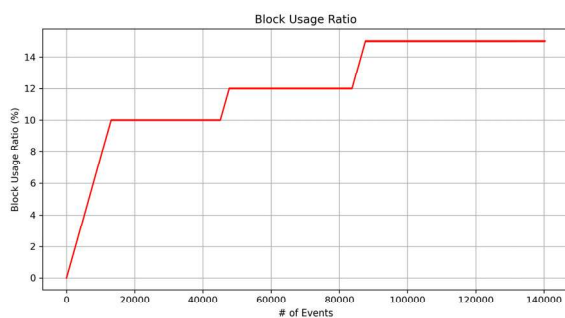


그림 5. Block Usage Ratio

그림 5는 실험을 통해 관찰된 block 사용 비율의 변화를 보여주며, 점진적인 garbage collection 전략 변경을 나타낸다. Garbage collection이 2000회 진행되기 전까지는 block 사용 비율이 0.1을 초과할 때마다 16개의 victim block을 선택하여 erase 작업을 진행하였다. 그 후, 2001회에서 3500회 사이의 garbage

collection에서는 비율이 0.12를 초과할 때마다 24개의 victim block을 선정하여 erase처리하였고, 3501회 이후에는 0.15를 초과하는 경우 32개의 victim block을 erase하였다.

IV. 결론 및 향후 연구 방향

본 연구에서는 disaggregated system에서 SSD의 write command 특성과 storage node의 SSD에서 생성되는 trace의 성질을 분석하였다. 또한 threshold 값을 기반으로 한 ‘Gradual Lazy Garbage Collection’을 적용하여, 해당 상황에서 block erase의 분포를 잘 관리할 수 있음을 확인하였다. 이 결과는 상당한 오버헤드를 가진 static wear leveling[6] 작업의 시점을 늦출 수 있다고 볼 수 있으므로 SSD의 효율적 사용과 수명 연장에 기여할 수 있음을 보여준다.

향후 연구에서 다양한 사용자 작업 시나리오를 반영한 trace 분석으로 확장하여, 장기적인 운용 시나리오에서의 wear leveling 불균일 문제를 해결하기 위한 방안을 모색할 수 있다. 이러한 연구를 통해 IaaS cloud system에서 SSD의 성능과 내구성을 최적화하기 위한 새로운 garbage collection 및 wear leveling 전략을 개발하는데 기여할 수 있을 것이다.

V. 사사

본 연구는 산업통상자원부(1415181094)와 KSRC 지원 사업인 미래반도체소자 원천기술개발사업(20019456)의 연구결과와 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구(No.2023-11-0830, AI 프로세서의 데이터 특성과 데이터 접근 특성에 최적화한 비휘발성 PIM용 메모리 모듈 및 메모리 컴파일러 개발)로 수행되었음.

참고문헌

- [1] S. S. Manvi and G. K. Shyam, "Resource management for Infrastructure as a Service (IaaS) in cloud computing: A survey," *Journal of Network and Computer Applications*, Vol. 41, pp. 424-440, 2014.
- [2] L. M. Vaquero, L. Roderó-Merino, and D. Morán, "Locking the sky: a survey on IaaS cloud security," *Computing*, Vol. 91, pp. 93-118, 2011.

- [3] Y. Kim, B. Tauras, A. Gupta, and B. Urgaonkar, "FlashSim: A Simulator for NAND Flash-Based Solid-State Drives," in Proc. of the First International Conference on Advances in System Simulation, pp. 125-131, Porto, Portugal, 2009.
- [4] A. Gupta, Y. Kim, and B. Urgaonkar, "DFTL: a flash translation layer employing demand-based selective caching of page-level address mappings," SIGPLAN Not., Vol. 44, No. 3, pp. 229-240, March 2009.
- [5] Red Hat, Technical Reference, https://access.redhat.com/documentation/ko-kr/red_hat_virtualization/4.3/html/technical_reference/qcow2.
- [6] Z. Jiao, J. Bhimani, and B. S. Kim, "Wear leveling in SSDs considered harmful," in Proc. of the 14th ACM Workshop on Hot Topics in Storage and File Systems, pp. 72-78, New York, NY, USA, 2022.