

SSD 요소별 소모전력 측정을 위한 다시점 측정방법

*유박은†, 원유집

한양대학교 전자컴퓨터통신공학과

e-mail : starhunter@ece.hanyang.ac.kr, yjwon@ece.hanyang.ac.kr

Measuring Power Dissipation for SSD

*Bal-Geun Yoo, Youjip Won

Dept. of Electrical and Computer Engineering
Hanyang University

Abstract

차세대 저장매체로 기대되는 Solid State Drive(SSD)는 저전력, 고성능이 그 장점으로 알려져 있다. 하지만 일반적인 인식과는 달리, 현재 판매되고 있는 SSD의 소모전력은 일반적인 Hard Disk Drive(HDD)와 비교해볼 때 에너지 절약의 관점에서 장점이 크지 않다. 이는 SSD의 제조사들이 성능적인 측면만을 강조하고 소모전력의 중요성을 간과한 결과이다. SSD의 소모전력을 줄이기 위해서는 현재 SSD의 소모전력을 정확히 분석하고 이를 토대로 모델을 세워야 한다. 본 논문은 이러한 연구의 첫 번째 단계로, SSD의 소모전력을 정확히 분석하는 방법을 제안한다. 거시적인 측정을 통해 타겟의 소모전력 패턴을 파악하고, 미시적인 측정을 통해 기존의 논문들이 놓치고 있는 소모전력의 피크 값을 올바르게 수집한다. 이러한 두가지 방법의 소모전력 측정을 통해서 다른 논문이 분석하지 못하였던 SSD 구성요소별 소모전력이 가능한 데이터를 수집하고, 소모전력 모델링의 토대를 제공한다.

I. 서론

SSD는 구동속도의 한계에 부딪힌 HDD를 대체하는 새로운 저장매체로 등장하였다. 또한 21세기의 그린IT 이슈와 맞물려 저전력 저장매체로서도 기대되고 있다. 일반적으로 SSD는 기계적 구동계를 가지고 있지 않기 때문에 HDD에 비해 소모전력이 적은 것으로 알려져 있다. 하지만 실제 측정결과는 그러하지 못하다. 그림 1은 3가지의 SSD에 대하여 컴퓨터의 전원을 켜고 SSD가 초기화 될 때 소모하는 전류를 측정하는 것이다. 실험에 따르면 SSD가 소모하는 전력은 제조사에 따라 매우 다르다. 이렇게 회사별로 SSD의 소모전력에 차이가 발생하는 것은 사용하는 부품의 차이일 수도 있지만, 내부에 사용하는 알고리즘, 즉 FTL이 아직까지는 소모전력을 고려하고 있지 못하기 때문이다.

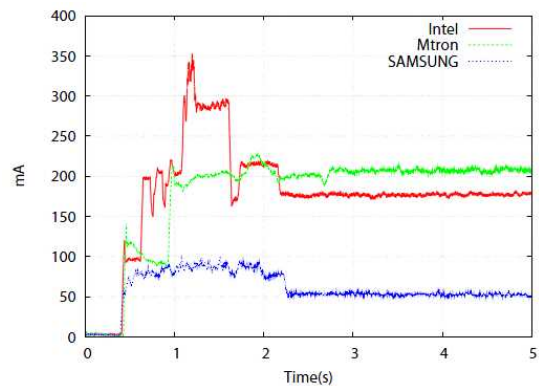


그림 1. 3가지 SSD의 Power On 소모전류

표 1. 3가지 SSD의 상세 스펙

	Mtron	Intel	Samsung
Model	SATA3525	X25-M	MPP
Interface	SATA2	SATA2	SATA2
Capacity	64GB	80GB	64GB
Nand Type	SLC	MLC	MLC
Max Read	100MB/s	250MB/s	80MB/s
Max Write	100MB/s	70MB/s	80MB/s
Channel	4ch	10ch	4ch
출시일	2008.1	2008.9	2008.11

앞으로 개발되고 출시될 새로운 SSD들은 더 큰 용량과 더 빠른 속도를 위해 지금보다 더 복잡한 내부 알고리즘을 사용하고, 병렬화 기법을 극대화해서 사용하게 될 것이다. 지금까지의 FTL 알고리즘을 그대로 적용하게 된다면 미래의 SSD들은 매우 높은 순간 소모전력 값을 가지게 될 것이 분명하다. 예상되는 문제점을 보완하기 위해서는 SSD의 구성요소를 고려한 소모전력 모델을 세우고, 이를 통해 저전력 FTL을 설계해야 한다. 이 논문은 그러한 과정의 첫 번째로, 실제 SSD에 대해 소모전력을 올바르게 측정하기 위한 측정방법을 제시한다. 그리고 실제 SSD에 적절한 워크로

† 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임 (No. ROA-2009-0083128)

드를 생성하는 개발된 프로그램을 사용하여 소모전류를 측정하였다. 2장에서는 관련연구를 설명하고, 3장에서는 측정하고자 하는 SSD의 구성요소를 설명하고, 4장에서는 SSD의 소모전력을 측정하는 실험방법론을 설명하고, 5장에서는 이 방법론에 따라 실험한 과정을 설명한다. 마지막으로 6장에서는 결론을 내린다.

II. 연구동기 및 관련연구

2.1 연구동기

연구의 초기에는 SSD의 성능분석이 주요 관심사였다. 반복적인 I/O를 수행할 때 SSD의 성능이 저하되는 현상을 관찰하기 위해 lseek을 사용하여 논리적으로 같은 위치에 반복적인 write를 수행하는 워크로드를 작성하고 다양한 SSD에서 실행하였다. SSD의 성능이 감소되는 패턴을 소모전류를 통해 분석해보고자 전류 probe를 이용하여 오실로스코프로 데이터를 수집하였다. 실험을 통해 소모전류의 패턴을 찾고자 하였고 그림 2의 실험 결과에서 Mtron SSD에서 주기적인 소모전류의 고점을 관찰할 수 있었다. 다른 SSD에서도 패턴을 찾고 소모전류를 분석하기 위해서 정형화된 워크로드가 필요하게 되었다.

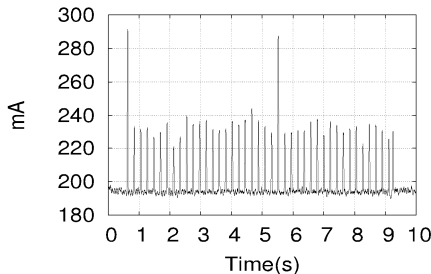


그림 2. Ext3 400MB 순차쓰기

2.2 관련연구

엔터프라이즈 시장에서는 디스크가 소모하는 전력이 문제가 되고 있다[1]. 문제의 해결을 위해서 기존 제품들의 소모전력을 측정하고 문제점을 파악하여야 한다. 디바이스의 소모전력 측정에는 직접적인 방법과 간접적인 방법이 있다. 직접적인 방법은 실제 디바이스를 하드웨어 측정 장비로 측정하는 방법이다. 간접적인 방법은 시뮬레이션 툴을 사용하는 방법과 소프트웨어적인 방법을 통해 결과 값을 추측하는 방법이 있다. HDD의 소모전력 측정을 위해 사용된 시뮬레이션 툴은 Dempsey [2]이 있다. Dempsey는 제조사에서 제공하는 HDD의 자세한 스펙 없이 자동으로 타겟 HDD로부터 정확한 성능과 소모전력에 대한 파라미터를 추출할 수 있다.

Shin et al.의 논문에서[3] SSD의 전력소모량을 좌우하는 요소로 플래시 셀의 종류, 디램 버퍼의 크기, 플래시칩의 병렬성, 펌웨어를 제시하였다. 소모전력의 측정은 SSD와 입력전원선 사이를 단선시켜, 그 사이에 0.5Ω의 저항을 연결하였다. 이 저항에 흐르는 전류를 Labview로 수집하고 관찰하였다. 실험에서 공장도 초기화 상태의 깨끗한 SSD를 Clean, 사용량이 누적된 SSD를 Dirty로 설정하고 Clean과 Dirty에 동일한 워크로드를 실행하고 소모전력을 비교하였다. 이 비교를 통해 Dirty에서 발생하는 추가적인 소모전력을 사용자

가 아닌 SSD 컨트롤러에서 실행하는 백그라운드 연산으로 가정하였다. 또한 입력 워크로드를 4kbyte부터 64kbyte까지 다양하게 입력하여 각 단계별 소모전력을 수집하였다. 하지만 이 방법은 입력 워크로드의 전체적인 동작 패턴만 파악이 가능하고 SSD의 내부 구성요소별로 소모하는 소모전력의 분석이 불가능하다.

Seo et al.의 논문에서는[4] SATA 전원선에 대하여 Signametrics SM2040 Digital PCI Multimeter를 사용하였고, DIO Tool이라 불리는 도구를 사용하여 순차쓰기, 임의쓰기, 순차쓰기, 임의쓰기의 워크로드를 입력하고, SSD의 소모전류를 측정하였다. 실험에 사용한 멀티미터는 데이터의 샘플링의 최소 단위가 1ms이다. 이 단위는 SSD 내부 구성요소의 동작시간을 요소별로 구분지어 분석하기에는 너무 길다. 이 방법 또한 SSD의 전체적인 동작 패턴만 파악이 가능하다. 내부 구성요소의 소모전력 파악을 위해서는 본 논문에서 제시하는 미시적 측정방법이 필요하다.

Park et al.의 논문에서는[5] SSD의 소모전력 모델링 요소를 channel과 way로 보고 하나의 way에 대한 읽기/쓰기 연산을 SSD 동작의 최소단위로 보았다. 실측을 바탕으로 구한 하나의 way에 대한 write를 기준으로 입력 워크로드의 용량에 따라 순차적으로 소모전류가 증가하는 모델링을 제안하였다. 실험에 사용한 SSD의 물리적 페이지 크기가 4kbyte 이지만 4kbyte의 순차쓰기 요청이 2개의 채널에 동시에 2kb씩 쓰이고 입력 워크로드의 크기가 클때 SSD의 최대 가용자원이 (channel과 way) 모두 동작한다는 모델을 사용하고 있다. 또한 하나의 way에 대한 write를 최소단위로 설정하고 예상 시나리오에 따라 최소단위의 소모전력이 단계적으로 더해진다고 가정한다. 본 논문에서는 물리적인 페이지 크기가 4kbyte인 SSD를 타겟으로 사용하고 있다. 실험결과에서 4kbyte의 소모전력과 8kbyte의 소모전력이 동일한것을 보여주고 있다. 이는 SSD의 내부 동작이 물리적인 페이지 크기별로 순차적으로 증가하지는 않는다는것을 보여준다.

III. SSD의 구조 및 측정요소

3.1 SSD의 내부 구조

일반적인 SSD의 내부구조는 그림 3와 같다.

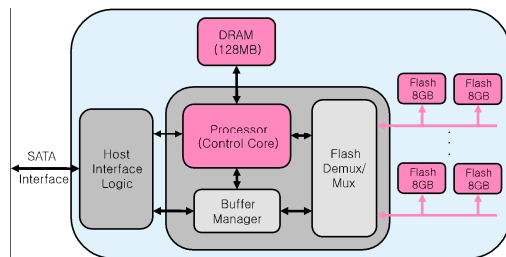


그림 3. 일반적인 SSD의 내부구조

SSD는 외부의 다양한 물리적인 연결을 위해 Host Interface Logic을 가지고 있다. 컨트롤러는 시스템 관점에서 SSD를 HDD와 동일하게 보이도록 변환하는 역할을 수행한다. 이를 Flash Translation Layer(FTL)이라고 한다. SSD 컨트롤러에는 내부 캐싱을 위한 DRAM과 디스크의 플래터에 해당하는 플래시칩이 있다. 플래시칩은 컨트롤러의 Demux/Mux에 연결되어

데이터를 동시에 병렬적으로 저장하고 로드 할 수 있다. 이를 SSD의 channel이라고 부른다. 하나의 channel에 연결된 다수의 플래시칩 구성을 way라 하고, 순차적으로 인터리빙 접근이 가능하다.

3.2 SSD 소모전력 측정요소

그림 3에서 설명한 SSD의 요소 중에서 본 논문에서 소모전력을 측정하고자 하는 구성요소는 5가지이다. 첫째는 SSD의 컨트롤러이고, 둘째는 SSD의 내부 DRAM이고, 셋째는 플래시칩이고, 넷째는 호스트 인터페이스이고, 마지막으로 내부 버스는이다. FTL의 핵심적인 역할을 수행하는 컨트롤러는 내부적으로 컨트롤 코어, 버퍼메니저, 플래시 Demux/Mux의 구성요소를 가지고 있다. 이 중 컨트롤 코어의 소모전력은 다른 두 구성요소에 비하여 상대적으로 매우 크다[6]. 그렇기 때문에 컨트롤 코어의 소모전력을 측정하는 전략은 곧 컨트롤러의 소모전력을 측정하는 전략이 된다.

IV. SSD 소모전력 측정 방법

본 연구에서 제안하는 방법은 정확한 측정을 위해 두가지 시점으로 소모전력을 측정하는 것이다. 먼저, SSD마다 각기 다른 특징을 가지고있기 때문에 거시적인 단위로 SSD의 소모전력 패턴을 측정해야한다. 이를 통해 타겟이 일정하게 전력을 소모하는지, 또는 일정 주기로 높은 전력을 소모하는지와 같은 소모전력의 패턴을 파악할 수 있다. 이러한 패턴의 파악을 위해 정형화된 워크로드를 제안한다. 하지만, 각 패턴별로 입력 워크로드에 따라 입력된 순간에 측정되는 개별적인 하나의 I/O 소모전력은 거시적인 측정방법으로는 측정이 불가능하다. 플래시메모리는 약 200us의 매우 짧은 프로그래밍 시간들[7] 가지고있다. 플래시 메모리의 프로그래밍 매커니즘은 매우 높은 전압을 플래시메모리에 인가하는 방법을 사용하고 있는데 이때 발생하는 순간적인 소모전력은 200us 보다 짧은 미시적인 데이터 샘플링 단위에서만 측정 가능하다. 데이터 샘플링 단위가 거시적일때는 가끔씩 수집되는 노이즈로 보이게 된다. 따라서 SSD의 소모전력을 측정하기 위해서는 거시적인 방법과 미시적인 방법이 모두 필요하다. 수집된 데이터는 적절한 표현을 위해 30개 간격으로 이동평균을 취하였다.

4.1 거시적 소모전력 측정

본 연구에서는 소모전력의 패턴 데이터 수집을 위해 write 연산에 관한 적절한 워크로드를 고안하였다. 부가적인 연산의 영향을 최소화 하기 위하여 파일시스템을 거치지 않고 Raw Device에서 실행하는 워크로드이다. 소모전력에 영향을 주는 요소는 I/O의 크기, I/O의 간격, I/O의 횟수 이다. I/O의 크기는 1kbyte 부터 2048kbyte까지 2의 배수로 증가시켰다. I/O의 간격은 500ms, 200ms, 100ms, 0ms의 4가지 간격을 두었다. I/O의 횟수는 100회 50회 10회로 하여 장기/단기 워크로드를 모두 관찰할 수 있게 하였다. 수집된 결과는 적절한 단위의 이동평균을 취했다. 그림 4는 동일한 I/O의 크기와 I/O의 횟수를 측정 간격을 달리하여 데이터를 수집하고 30단위로 이동평균을 적용한 결과이다. 데이터 수집 시간이 길어지면 수집되는 소모전력 최대값이 누락되고 있다는 것을 알 수 있다. 그렇다면

이 방법은 SSD 소모전력의 패턴을 분석하기에는 용이하지만 정확한 소모전력 데이터를 수집하지는 못한다는 것을 알 수 있다. 이는 데이터의 수집 간격이 플래시 메모리의 프로그래밍 시간보다 길기 때문이다. 이를 해결하기 위해 새로운 데이터 수집방법이 필요하게 되었다.

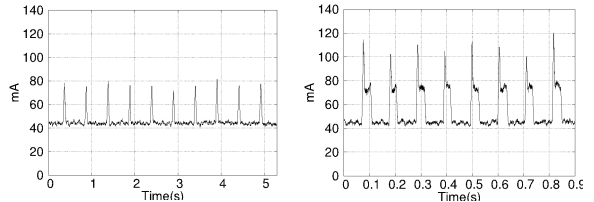


그림 4. 자료수집 간격이 다른 2가지의 4K 쓰기

4.2 미시적 소모전력 측정

일반적으로 플래시 메모리의 프로그래밍 시간은 200us[7]이다. 섹션 4.2에서 사용한 워크로드에서 I/O의 간격을 100ms로, I/O의 횟수를 100회로 고정하고 I/O의 크기를 달리하고 오실로스코프에서 I/O의 소모전력을 실시간으로 수집하였다. 그림 5를 보면 그림 4에서 측정한 최대 소모전력값 보다 높은 것을 알 수 있다. 플래시 메모리의 프로그래밍 시간보다 샘플링 간격이 길다면, 소모전력의 최대값을 올바르게 측정하지 못한다. 정확한 소모전력 데이터 수집을 위해서는 사용하는 플래시메모리의 프로그래밍 시간보다 데이터의 수집간격을 짧게 하여야만 한다. 실험에서 사용한 Tektronix TDS3032에서 데이터 수집 간격을 10us로 설정하고 실험하였다. 그림 5에서 4kbyte와 8kbyte, 16kbyte의 소모전류는 동일한 패턴과 동일한 피크값을 가지고 있다. 이것은 타겟 SSD가 용량이 다른 3가지 워크로드에 대하여 동일한 전류를 소모한것을 알 수 있다. 입력 워크로드의 용량이 32kbyte가 될 때 소모전류의 피크값이 하나 증가하였고, 64kbyte 이상이 될 때 소모전류의 피크값과 시간이 변하고 있다. 1024kbyte 순차쓰기 연산은 동일한 512kbyte의 반복이다. 이러한 결과는 Park et al.의 논문[5] 결과와는 매우 다르다. 이러한 결과의 검증을 위해 타겟을 다양화하여 추가적인 실험이 필요하다. 그림 5의 (d)(e)(f)는 이동평균을 취하지 않은 샘플링 데이터이다. 실제로 소모되는 전류의 피크값은 미시적 관점으로 이동평균을 취하지 않았을때 측정 가능하다.

V. 실험 및 결과

5.1 실험환경

실험환경은 SSD에 입력되는 5V 전원 선에 대해 전류 Probe를 통해 오실로스코프로 소모전류를 관찰하였다. 특정 워크로드를 SSD에 대해 실행하고, 해당 워크로드의 소모전류 데이터를 오실로스코프에서 수집하였다. 오실로 스코프는 Tektronix사의 TDS3032 모델을 사용하였다. 실험에 사용된 SSD의 모델명은 MMCRE28G5MXP-0VB 이고 S-ATA2 인터페이스를 사용하며, MLC 타입의 128GB의 용량을 가지고 있으며, 내부에 8GB용량의 플래시 패키지를 16개가 8채널로 연결되어 있다. 내부 DRAM Cache는 128MB 이다.

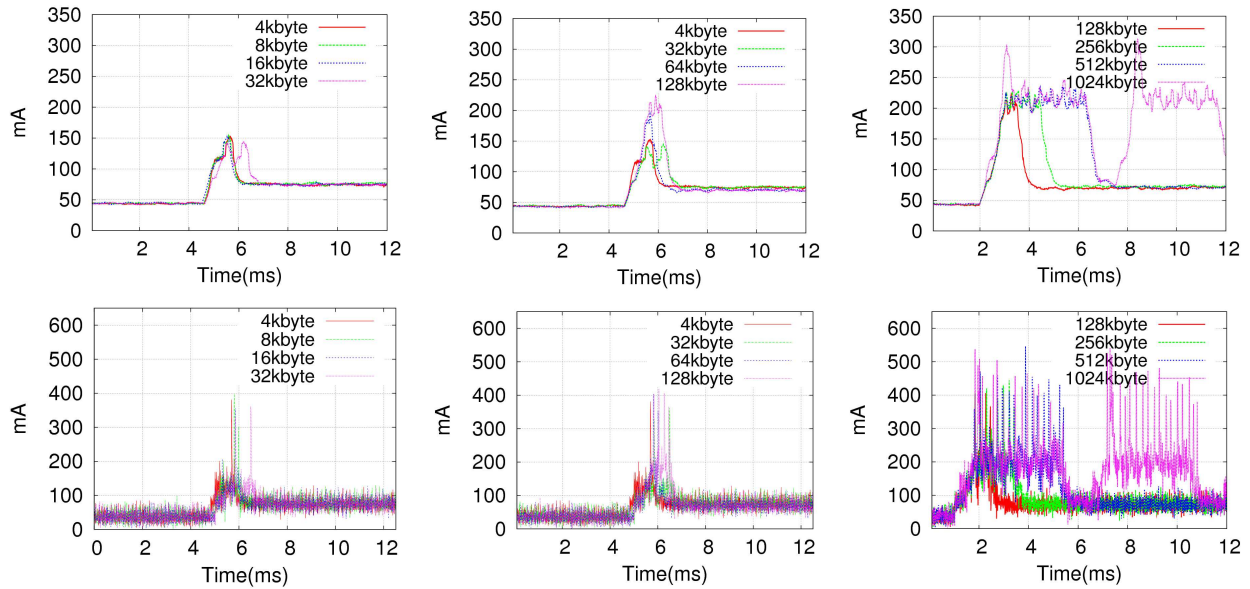


그림 5. 미시적 소모전력 측정방법으로 수집한 소모전력 데이터

실험에 사용된 전류 probe는 회로의 단선 없이 측정하고자 하는 전원선의 피복 위에 연결하는 장비이다. 타겟 SSD의 전원 입력 선에 전류 probe를 연결하고 이를 오실로스코프에서 관찰하였다. 수집된 데이터는 노이즈 제거를 위해 30개의 단위로 이동평균을 취하였다. 이는 피크값의 데이터를 누락시키지 않으면서 노이즈만을 제거할 수 있는 최소 단위를 휴리스틱하게 선정한 값이다. 실험에는 Linux Kernel 2.6.24.4 버전에서 Ext3 파일시스템을 사용하였다. 실험의 기본 설정으로 look-ahead와 read-ahead 기능을 off 하였고, 상황에 맞게 hdparm 명령으로 타겟 디바이스의 write-caching 기능을 on/off 하였다. 4장에서 소개한 소모전력 측정방법에 맞게 10us 간격으로 데이터를 수집하였다.

5.2 실험결과

그림 5을 보면 입력 워크로드가 일정 용량 이상이 될 때 소모전류 값에 변화가 생기는 것을 볼 수 있다. 이는 타겟 SSD가 원자적으로 가용자원을 구동하는 것이 아니라, 필요 없는 자원을 저용량의 워크로드 입력 시에도 필요한 부분 이상의 가용자원을 구동시키고 있다는 것을 알 수 있다. 또한 수집 간격이 클 때 수집하지 못했던 소모전력 데이터들을 모두 수집하고 있다는 것을 알 수 있다. 이러한 정확한 자료 수집을 통해 정확한 자료 분석이 가능하다.

VI. 결론

이 논문은 SSD의 소모전력 분석에 앞서 정확한 자료수집의 방법을 제시하고 있다. 거시적 단위로 소모전력을 측정함으로써 타겟의 소모전력 패턴을 알 수 있고, 기존의 연구들에서 고려하지 못하고 놓치고 있는 데이터들에 대해서 플래시 메모리의 프로그래밍 시간보다 짧은 단위인 미시적 단위로 데이터를 샘플링함으로써 누락되는 데이터 없이 자료수집이 가능하다.

4장에서 소개한 측정 방법을 다양한 SSD에 적용하여 데이터를 수집하고 분석한다면, 현재 출시된 SSD들의 소모전력에 대한 정확한 분석과 모델링이 가능할 것이다. 앞으로 본 논문의 방법을 토대로 SSD의 주요 구성요소들의 소모전력에 대하여 수식적인 모델링이 필요하다.

참고문헌

- [1] S. Lee, B. Moon, and C. Park, "Advances in flash memory ssd technology for enterprise database applications," in Proceedings of the 35th SIGMOD international conference on Management of data. ACM, 2009, pp. 863 - -870.
- [2] J. Zedlewski, S. Sobti, N. Garg, F. Zheng, A. Krishnamurthy, and R. Wang, "Modeling hard-disk power consumption," in FAST '03: Proceedings of the 2nd USENIX Conference on File and Storage Technologies. Berkeley, CA, USA: USENIX Association, 2003, pp. 217 - -230.
- [3] D. Shin, "Power Analysis for Flash Memory SSD," in NVRAMOS, 2010, Spring.
- [4] E. S. S. Y. P. B. Uргаonkar, "Empirical analysis on energy efficiency of flash-based ssds," in HotPower'08: Workshop on Power Aware Computing and Systems. USENIX Association, 2008.
- [5] J. Park, S. Yoo, S. Lee, and C. Park, "Power Modeling of Solid State Disk for Dynamic Power Management Policy Design in Embedded Systems." Springer, 2009, pp. 24 - -35.
- [6] SAMSUNG. Electronics, "S3C49F9X SOLID DISK CONTROLLER Data Sheet."
- [7] SAMSUNG. Electronics, "K9XXG08UXA Flash Memory Data Sheet," 2006.