

2020 한국차세대컴퓨팅학회 하계학술대회



장 소 : 제주 JDC(엘리트빌딩)
일 시 : 2020. 8. 20(목) 14:00 ~ 8. 22(토) 12:00
주최 · 주관 : 한국차세대컴퓨팅학회, 제주드론산업협회

스토리지 데이터 델타 측정을 통한 시스템 에이징 측정 도구 개발 및 상관관계 분석

System Aging Measurement Tool Development And Correlation Analysis Through Storage Data Delta Measurement

장은태¹, 최상훈¹, 박기웅^{2,*}

Eun-Tae Jang, Sang-Hoon Choi, Ki-Woong Park

¹(05006) 서울특별시 광진구 능동로, 세종대학교 시스템보안연구실

²(05006) 서울특별시 광진구 능동로, 세종대학교 정보보호학과

euntaejang@gmail.com, csh0052@gmail.com, woongbak@sejong.ac.kr

요 약

운영체제에서는 사용자에게 의해 소프트웨어의 설치와 제거가 반복되며 시스템을 오랜 기간 동안 사용하게 되면 성능이 저하되는 현상이 자주 발견되는데, 이러한 현상을 시스템 에이징이라고 한다. 운영체제에서는 수많은 소프트웨어들이 사용자의 필요에 의해 설치되고 제거되고 있지만 운영체제는 설치된 소프트웨어의 완전 제거를 보장하지 못하고 있다. 이러한 문제점으로 인해 사용자가 운영체제에 설치된 소프트웨어를 제거할 경우 설치 파일의 일부 혹은 소프트웨어가 사용했던 로그파일, 임시파일 등을 시스템에 남겨두고 제거되는 것을 실험을 통해 확인할 수 있었다. 본 논문에서는 운영체제에 설치된 소프트웨어가 제거 될 경우 깨끗하게 지워지지 않고 수많은 파일들을 하드디스크에 남기고 제거되는 것을 실험을 통해 확인하며 남아있는 파일의 개수와 크기를 측정할 수 있는 도구를 개발한다.

키워드: 시스템 에이징, 완전 제거, 도구 개발, 소프트웨어, 하드디스크

1. 서론

운영체제는 사용자가 컴퓨터를 편하게 사용할 수 있도록 인터페이스 환경을 제공해주며, 컴퓨터 시스템의 자원을 효율적으로 관리해주는 역할을 수행한다. 하지만 운영체제는 사용 할수록 성능이 저하되는 노화현상이 발생하고 있으며 이러한 현상을 시스템 에이징[1]이라고 한다. 하지만 운영체제의 성능저하를 유발하는 요소는 많으며, 정확한 성능저하의 원인을 찾는 것은 매우 어렵다[2]. 논문에서는 운영체제가 파일을 제거할 때 깨끗하게 지우지 못하는 것을 실험을 통해 확인 하였고, 이러한 문제점이 운영체제의 성능저하를 유발할 수 있는지

확인한다. 운영체제에서는 수많은 소프트웨어들이 사용자에게 의해 설치되고 제거되고 있지만 운영체제는 설치된 소프트웨어가 제거될 경우 완전 제거를 보장하지 못하는 문제점을 가지고 있다. 소프트웨어가 운영체제에서 제거 될 경우 소프트웨어는 많은 파일들을 하드디스크에 남기게 되며 이러한 제거 되지 않은 파일들은 소프트웨어가 설치 될 때 설치 파일의 일부 또는 소프트웨어가 사용했던 임시파일, 로그파일, 링크파일 등이 있었다. 운영체제에서는 컴퓨터 내 많은 파일들을 손쉽게 검색하기 위해서 자동으로 파일들을 색인해주는 프로세스가 존재한다(ex. 윈도우즈 운영체제의 searchIndexer.exe). 이러한 프로세스는 사용자가

* 교신저자

운영체제 내의 파일을 빠르게 검색할 수 있도록 도와주지만 운영체제 내의 더미 파일, 즉 필요 없는 파일이 많아질수록 CPU의 점유율이 높아져 시스템의 성능을 저하시킨다. 본 논문에서는 국내 소프트웨어 중 사용자가 많이 설치한 소프트웨어 4개를 선택하여 실험을 진행하였다. 운영체제 중 가장 높은 점유율을 차지하는 윈도우즈 운영체제에서 소프트웨어가 설치되고 제거 될 때 얼마나 많은 파일들을 남기고 제거 되는 지 확인하기 위해, 크게 세 단계로 나누어 실험을 진행하였다. 실험을 위해 논문에서는 측정 도구를 개발하여 각각의 단계의 하드디스크를 덤프하고, 디스크 덤프파일에서 파일리스트를 추출하였으며, 추출한 파일들을 비교하여 소프트웨어 제거 후에 얼마나 많은 파일을 운영체제에 남기고 제거 되었는지 확인하였다.

본 논문의 구성은 다음과 같다. 2장에서는 운영체제 및 응용프로그램을 대상으로 시스템 에이징의 원인을 연구들을 정리하고, 3장에서는 소프트웨어가 운영체제에서 완벽하게 제거 되지 않는 것을 실험을 통해 증명하며, 이러한 문제점이 시스템의 성능에 영향을 미칠 수 있음을 보인다. 4장에서는 결론 및 추후연구 방향을 제시한다.

2. 관련 연구

시스템 에이징 현상의 원인을 연구하기 위하여 연구자들은 운영체제에서 부터 응용프로그램에 이르기까지 다양한 연구를 수행하였다.

- Sachin Garg과 Aad van Moorsel은 소프트웨어 에이징은 운영체제의 메모리 스왑 공간 소진, 운영체제의 리소스 사용량 증가가 소프트웨어 에이징의 핵심적인 원인이 되며[1], 소프트웨어에 영향을 주는 리소스에 대해 모니터링을 수행할 중요한 자원을 식별하는 것이 우선시 되어야 한다고 강조하였다.
- Domenico Cotroneo은 오랜 시간 동안 지속적으로 실행되는 소프트웨어는 시간에 지남에 따라 성능이 저하되며 오류발생률이 증가하기 때문에 실험을 위해 커널 추적도구를 사용하여, 여러 커널 하위 시스템의 관련 매개 변수를 수집하고 통계 분석을 통해 리눅스 운영체제의 에이징 원인을 확인하였다[2]. 실험 결과 리눅스에서 에이징의 원인은 Process Management, File system이 핵심 원인임을 도출하였다.

- Luis Silva은 JVM위에서 실행되는 SOAP 기반 서버의 소프트웨어 에이징을 분석하였다. Luis Silva은 소프트웨어 에이징의 심각한 문제를 해결하기 위해 SLA 기반의 서버 통신 계약을 만들어 프로그램의 지속적인 성능 수준을 유지하는 Rejuvenation 기술[3]을 제안하였다.
- Kenichi Kourai은 가상 머신(VM)을 관리하는 가상 머신 모니터(VMM) 소프트웨어 에이징에 대해 연구하였으며[4] VMM의 성능 저하 또는 충돌 관련 에러는 모든 VM에 영향을 주기 때문에, VMM의 에이징을 막기위하여 VMM Rejuvenation 기술인 warm-VM을 제안하였다. 해당 기술을 사용하면 메모리 이미지에 액세스하지 않고 VM을 일시 중단하고 빠르게 다시 시작하여 VM을 재부팅 하지 않아도 VMM만 효율적으로 재부팅 하여 VMM을 재부팅하는 동안 메모리 이미지를 유지한다.
- Michael Grottk은 장시간 실행되는 소프트웨어 시스템의 고장률이 증가하거나 성능이 점진적으로 저하되는 현상이 많이 목격되며, 이러한 현상은 내부 오류의 누적 및 운영체제의 리소스 고갈이 원인이 되는 소프트웨어 에이징 현상이라고 하였다[5]. 이를 막기 위해서는 시스템 리소스 사용량의 측정, 분석, 예측이 필요하며, Michael Grottk는 웹 서버 안에 리소스 사용 단계를 관찰하여 시스템 리소스 사용량 및 매개변수에 대한 데이터를 수집하였다. 수집 된 데이터를 바탕으로 통계 분석을 통해 향후 리소스 사용을 적절하게 예측하여 에이징을 사전 예방하는Software Rejuvenation 기법을 제안하였다.

3. 프레임워크 개발 및 실험

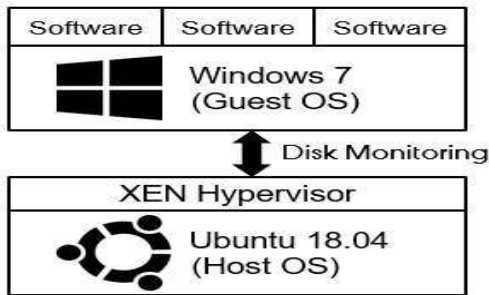
해당 장에서는 소프트웨어가 제거 될 경우, 설치 프로그램이 완벽하게 지워지지 않고 많은 더미 파일들을 하드 디스크에 남기고 제거되는 것을 실험을 통해 도출한다. 또한 실험을 위해 도구를 개발[6]하여 디스크 덤프, 파일 추출, 파일 비교를 수행하여 소프트웨어가 제거 된 상태와 소프트웨어가 설치 된 상태를 비교했을 때 어느 정도의 파일이 디스크에 남겨졌는지를 확인한다. 이를 위해 네이버 소프트웨어 사이트의 순위에 따라국내 소프트웨어 중 사용자들이 많이 설치하는 소프트웨어 3개를 선택하여 실험을 진행하였다.

- KM Player
- 알약

- V3
- nProtect

3.1 실험 환경

소프트웨어가 하드디스크에 남기고 제거 되는 파일들의 정보를 수집하기 위해 오픈소스 파이썬 라이브러리인 sleuthkit를 활용하여 도구를 개발하였다. 해당 라이브러리를 사용하면 디스크 덤프파일에서 파일의 리스트를 추출할 수 있다. 디스크 덤프 및 추출한 파일의 리스트를 분석하기 위해 XEN 가상환경을 사용하였다. 가상환경을 사용하지 않을 경우, 디스크를 덤프를 수행하고 덤프파일에서 파일을 추출하는 작업에서, 많은 오버헤드가 발생하여 시스템의 성능이 저하되는 것을 확인할 수 있었기 때문에 가상환경으로 실험 환경을 <그림 1>과 같이 구축하였다.

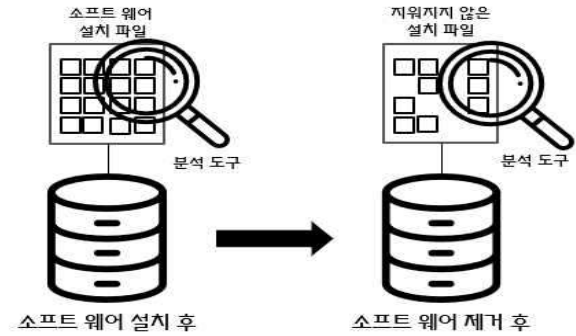


(그림 1) 파일리스트 추출을 위한 환경구축

3.2 파일리스트 분석을 위한 프레임워크 개발

소프트웨어가 제거 될 경우 하드디스크에 지워지지 않고 남아있는 파일들을 확인하기 위해 크게 3가지의 상태로 디스크를 덤프하여 실험을 진행하였다.

- Pure State : 윈도우즈 운영체제에 아무것도 설치하지 않은 초기 상태
- Installed State: 윈도우즈 운영체제에 사용자가 소프트웨어 하나를 설치한 상태
- Removed State: 윈도우즈 운영체제에서 사용자가 설치한 소프트웨어를 제거한 상태



(그림 2) 파일리스트 분석을 위한 프레임워크

세 가지 상태에서 디스크를 덤프 한 뒤 덤프파일에서 파일의 리스트를 추출하고 이를 비교하여 소프트웨어가 제거 된 후에 남기고 간 파일의 개수와 크기를 확인한다. 먼저 소프트웨어 설치 파일들을 추출하기 위해 Installed State에서 추출한 파일과 Pure State에서 추출한 파일사이의 차집합 연산을 통해 소프트웨어 설치 파일들을 추출하였다. 그 다음 추출한 소프트웨어 설치 파일 리스트와 Removed State에서 추출한 파일의 리스트의 공통된 파일, 즉 교집합 연산을 사용하여 소프트웨어 제거 후에 디스크 남아 있는 설치 파일들의 개수와 크기를 알아 낼 수 있었다. 해당 과정은 다음과 같다.

$$\begin{aligned} \text{Software Installed File} &= \text{Installed State} - \text{Pure State} \\ \text{Not Removed File} &= \text{Software Installed File} \cap \text{Removed State} \end{aligned}$$

3.3 실험 결과

선정한 4개의 소프트웨어를 대상으로 실험을 진행하였을 때 각각의 소프트웨어에 대해 제거 되지 않고 디스크에 남겨져 있는 파일들의 개수와 크기를 도출할 수 있었다. 실험 결과는 <표 1>과 같다. 제거 되지 않은 파일들은 해당 소프트웨어가 사용했던 로그파일, 임시파일 및 설치 파일의 일부가 있었다. 해당 실험 결과는 약 5%의 오차가 있는 것을 확인할 수 있었다.

<표 1> 소프트웨어 제거 후의 파일의 개수와 크기

소프트웨어	지워지지 않은 파일의 개수/ 설치된 파일 개수	지워지지 않은 파일의 크기/ 설치된 파일 크기
KM Player	81.2864775997	71.8020315782%
알약	31.048591049%	63.0084538197%
V3	62.632696390%	54.5364154390%
nProtect	78.9366053170%	58.4887772091%

[2]Software Aging Analysis of the Linux Operating System
 [3]Software Aging and Rejuvenation in a SOAP-based Server
 [4]A Fast Rejuvenation Technique for Server Consolidation with Virtual Machines
 [5]Analysis of Software Aging in a Web Server
 [6]https://github.com/zj3t/kingpc/blob/master/solve_r.py

4. 결론 및 추후연구

본 논문에서는 실험을 통해 운영체제에 소프트웨어가 설치되고 제거될 경우 운영체제는 설치된 소프트웨어에 대한 완전 제거를 보장하지 못한다는 것을 증명하였다. 국내 소프트웨어들 중 사용자가 많이 설치하는 4개의 소프트웨어를 대상으로 실험한 결과 소프트웨어가 설치되고 제거 될 때 많은 파일들을 디스크에 남기고 제거 되는 것을 실험을 통해 확인하였으며 디스크에 제거 되지 않고 남겨진 파일들은 거의 설치 프로그램의 절반 이상의 크기를 차지하고 있었다. 운영체제에서는 컴퓨터 내 많은 파일들을 손쉽게 검색하기 위해서 자동으로 파일들을 색인해주는 프로세스가 존재한다. 해당 프로세스는 디스크 내 더미 파일이 많아질수록 CPU의 오버헤드가 증가하게 되고 결과적으로 시스템의 성능을 저하시키게 된다. 추후 연구에서는 윈도우즈 운영체제뿐만 아니라 리눅스 등의 다른 운영체제에서도 소프트웨어가 제거 될 때 많은 더미파일들을 남기고 가는지 확인하며 소프트웨어가 제거 될 때 디스크에 남기는 파일들이 시간이 지남에 따라 시스템에 오버헤드를 축적시킬 수 있음을 확인한다.

Acknowledgement

본 연구는 2019년도 과학기술정보통신부의 재원으로 정보통신 기획평가원의 지원(No. 2019-0-00273) 및 한국연구재단 연구과제(NRF-2020R1A2C4002737)의 지원을 받아 수행된 연구임.

참고문헌

[1]A Methodology for Detection and Estimation of Software Aging