효과적인 로그 분석을 위한 로그분석기 설계 및 구현

우연옥[O], 황성철, 이상인, 강흥식
인제대학교 컴퓨터공학과
{poppi99[O], mslove1}@nate.com, delispai@yahoo.co.kr, hskang@nice.inje.ac.kr

A Design and Implementation Linux Log Analysis System for effective Log Analysis

Yeaonok Woo[O], Sungchul Hwang, Sangin Lee, Heungseek Kang
Dept. of Computer Engineering, Inje university

요 약

리눅스에서는 시스템상의 문제와 보안상의 이유로 특정 파일에 로그를 기록하게 된다. 이 로그파일을 보고 시스템의 문제를 진단하고 시스템을 효과적으로 관리할 수 있다. 또한 이것은 시스템을 안전하게 지키기 위한 도구가 될 수 있다. 그러나 이러한 로그파일은 항상 백업이 필요한 방대한 양의 로그를 지니고 있어 많은 양의 디스크 공간을 차지하고 있으며, 혹시 무슨 일이 있을지 몰라 정기적으로 남기고는 있는데 무슨 내용이 담겨 있는지 의미를 제대로 이해하지 못해 별 도움이 안될 뿐만 아니라, 분석하기도 쉽지 않은 어려움이 있다. 이에 본 논문에서는 보다 효과적으로 리눅스 시스템의 로그파일을 분석가능하며, 전문가가 아닌 초급 시스템 관리자들도 충분히 이해할 수 있는 리눅스 시스템 로그 분석기를 설계 및 구현해 보았다.

1. 서 론

집에 절도범이 침입하였을 경우, 운손잡이, 창문 등에 난 지문이나 거실바닥의 발자국, 또는 범인의 머리카락 등을 수집하여 분석한다. 범인이 남긴 이러한 사소한 단서들은 범인을 추적하는데 중요한 자료가 된다.

컴퓨터 시스템을 불법적으로 침입한 공격자들도 시스템 여기저기에 이러한 단서들을 남긴다. 불법 침입자의 침입흔적은 시스템의 각종 로그파일에 남는다. 시스템에 대한 스캔 행위, exploit 툴을 이용한 공격, 특정 사용자 계정으로의 접속, root 권한의 획득, 트로이목마 설치, 자료 유출 및 삭제 등 공격자의 행위 모두가 시스템에 의해 감시되고 로그로 남게된다.

리눅스 시스템에서는 각종 데몬과 커널에서 매일 엄청난 로그를 남기고 있다. 하지만 이러한 로그의 의미를 제대로 이해하지 못하는 시스템 관리자에게 이러한 로그는 디스크만 잡아먹는 쓸모없는 존재일 수밖에 없을 것이다. 이에 본 논문에서는 초보 사용자도 쉽게 원격에서 로그파일을 분석하고 관리할 수 있는 로그분석기를 설계하고 구현하였다.

2. 관련 연구

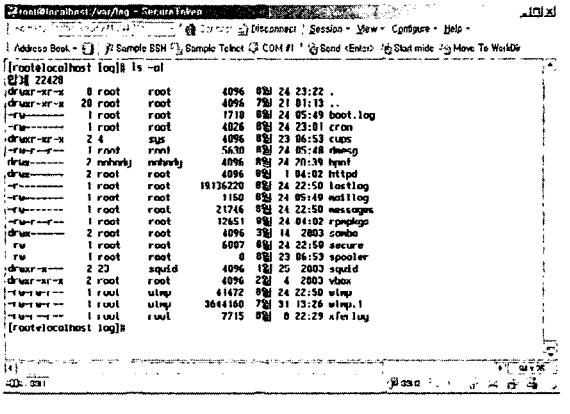
2.1 리눅스 로그

로그란 흔적이라고도 하며 운영체제나 응용 프로그램이 가동될 때 일어나는 이벤트를 기록한 파일을 의미한다. 로그를 남기는 작업을 로깅이라고 하며, 로그를 남김으로써 시스템 관련 운영 정보를 저장할 수 있게 된다. 관리자는 이것을 통해서 시스템에 대한 일련의 사항들을 모니터링 할 수 있다.

따라서 시스템관리에 있어서 정기적으로 확인 및 점검해야할 사항들 중에 가장 중요한 것이 로그파일들이다. 시스템에 이상이 있거나 보안의 위험을 감지하기 위해서는 시스템에서 남겨지는 메시지를 확인해야 한다. 모든 시스템에는 특정작업이 발생한 후에는 반드시 로그가 남겨지며, 관리자는 이를 정기적으로 점검을 해야한다. [1]

2.2 리눅스 로그파일

리눅스에서는 /var/log 디렉토리에서 시스템의 모든 로그를 기록하고 관리한다. 시스템의 /etc/syslog.conf 파일에서 시스템 로그파일들의 위치를 지정하고 있다. [2]



[그림 1] /var/log 디렉토리

2.2.1 utmp, utmpx

시스템이 현재 로그인한 사용자들에 대한 상태를 가지고 있다. /var/run/utmp 파일에 바이너리 형태로 저장되어 vi 등의 편집기로는 확인할 수 없다. 따라서 who, w, whodo, users, finger 등의 명령어가 utmp 파일을 참조하여 관련 정보를 사용자가 볼 수 있는 형태로 보여준다.

2.2.2 wtmp, wtmpx

사용자들의 로그인, 로그아웃 정보를 가지고 있다. utmp 파일과 마찬가지고 바이너리 형태이며 last 명령을 이용하여 내용을 확인 할 수 있다. utmp 파일이 현재 로그인해 있는 사용자에 대한 snapshot이라고 하면, wtmp는 지금까지 사용자들의 로그인, 로그아웃 히스토리를 모두 가지고 있다. 시스템의 shutdown, booting 히스토리까지 포함하고 있어, 해킹 피해시스템 분석에서 대단히 중요한 로그이다.

2.2.3 secure

보안과 관련된 주요한 로그를 남기며, 사용자 인증 관련된 로그를 포함하고 있다. 로깅 데몬인 syslog 데몬에 의해 남겨지는데 바이너리 파일이 아니므로 vi 등의 편집기로도 확인 가능하다.

2.2.4 lastlog

각 사용자가 가장 최근에 로그인 한 시간이 기록되는 파일로서 사용자가 시스템에 로그인 할 때마다 기록된다. 동일한 사용자에 대해서는 이전 내용을 overwrite 함으로써 갱신한다.

2.2.5 sulog

su(substitute user) 명령어를 사용한 결과가 저장되는 파일이다.

2.2.6 messages

콘솔 상의 화면에 출력되는 메시지들은 messages 로그파일에 저장이 된다. 대단히 방대한 정보를 포함하고 있어 시스템 관리자가 시스템 장애 원인을 찾아내기 위해서도 messages 파일을 점검한다. 이 파일에는 파일시스템 full, device failure, 시스템 설정 오류 등의 다양한 내용을 가지고 있다.

3. 로그 분석기

3.1 로그분석의 필요성

시스템 관리자가 로그파일을 분석함으로써 시스템 전반에 대한 상태를 파악할 수 있으며 문제 발생 시 원인을 파악하고 대처할 수 있게 된다. 특히 로그는 공격자의 공격으로부터 일어나는 행동을 기록함으로써 공격의 증거 자료로 활용할 수 있기 때문에 관리를 철저히 해주어야 한다. 또한 정기적으로 로그를 살펴봄으로써 보안을 높여야 한다.

3.2 현재 로그분석의 문제점

각종 데몬과 커널에서는 매일 엄청난 로그를 남기는데 대부분의 로그분석 작업은 수동으로 해야하는 번거로움이 따른다. 또한 일반적인 사용자나 초보적 수준의 시스템 관리자에게는 로그의 의미를 제대로 이해하고 분석하기에는 많은 어려움과 수고를 필요로 한다.

3.3 현재 로그분석기

3.3.1 log-watcher

해당 로그파일을 하루 단위로 종합적으로 정리를 해서 메일로 보내주는 데몬이다. [3]

3.3.2 log-scanner

Perl로 작성되어 있으며, TCP wrapper 패키지와 결합되어 사용할 수 있도록 설계되었다. 로그 스캐너의 목적은 시스템 관리자가 로그파일에서 미리 정의한 이상 행위를 발견했을 때 관리자에게 연락하도록 로그파일을 설정하는 것이다. [4]

3.3.3 swatch

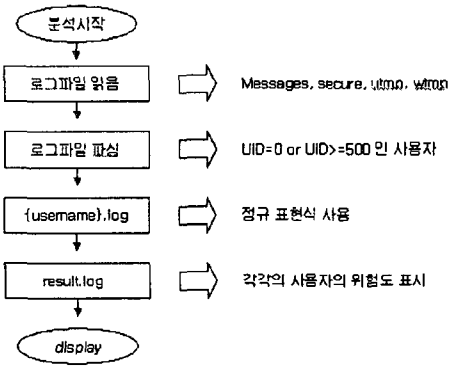
Perl로 작성된 실시간 모니터링 툴이다. 로그를 모니터하면서 특정 패턴에 반응해 요청한 작업들을 해 줄 수 있다. [5]

4. 리눅스 로그 분석기(Linux Log Analysis System)

4.1 LLAS의 설계

LLAS는 초보 사용자들도 방대하고 어려운 리눅스 로그를 쉽고 간편하게 원격에서도 분석할 수 있도록 설계하였다. 먼저 리눅스의 중요 로그파일들을 읽어들인 다음 파싱을 한다. 파싱은 UID를 우선으로 하며 UID가 0(root)이거나 500이상인 사용자에 한해서만 추출하는데 그 이유는 UID 1번부터 499까지는 특수 사용자들이기 때문이다. 이렇게 추출된 사용자들에 대한 로그들을 {username}.log 파일로 저장한다.

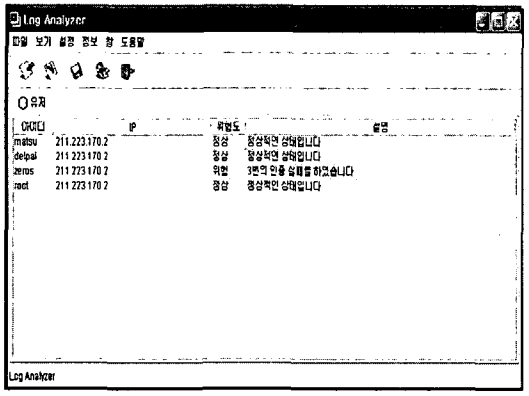
그 후 각 사용자들에 대한 로그파일을 한 줄씩 읽어들여 정규표현식을 사용하여 SSHD, LOGIN, FTP, SU, PASSWD 서비스들의 Session Opened, Session Closed, Authentication Failure, Invalid Users, Unknown Entries등에 대한 통계를 내어 그에 대한 결과로 result.log 파일을 만든다. 이 파일은 각각의 사용자들에 대한 위험도를 알려주기 위해 사용되어진다. 다음 [그림2]는 LLAS의 설계에 대한 순서도이다.



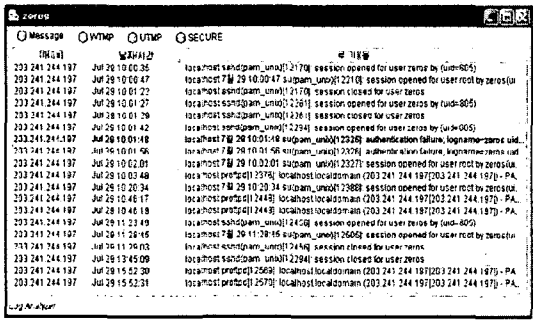
[그림 2] LLAS 순서도

4.2 LLAS의 구현

LLAS는 국내에는 잘 알려지지 않았지만 생산성과 확장성이 뛰어난 언어인 python과 객체지향언어인 Java로 작성되었으며 Linux 9.0 환경과 Windows 2000 Professional환경에서 실험하였다. [그림3, 4]는 LLAS의 실행화면이다. [6][7][8]



[그림 3] LLAS 분석화면 1



[그림 4] LLAS 분석화면 2

5. 결론 및 향후과제

본 논문에서는 리눅스 시스템의 로그 파일에 대한 분석을 위한 시스템인 LLAS를 설계 및 구현해 보았다. 리눅스 시스템의 로그파일에 대한 중요성은 이미 별다른 언급이 없어도 많은 사용자들이 인지하고 있는 부분이다. 하지만 리눅스 시스템의 로그 파일의 관리는 많은 전문지식 없이는 해결하기 어려운 과정이다. 이유는 각각의 로그파일들이 가지고 있는 의미가 다르기 때문이고, 또한 이러한 로그 파일에 대한 백업이나 해석 역시 초급 관리자에게는 상당히 힘겨운 부분들이었다. 따라서 특정 관리자가 아닌 초급 관리자 혹은 리눅스 시스템에 대한 많은 지식이 없는 사용자라도 손쉽게 원격에서 리눅스 서버를 분석할 수 있는 LLAS를 설계하고 구현하였다. LLAS는 특징은 기존의 많은 로그분석기처럼 터미널 화면에서의 레이아웃에서 벗어나 GUI형태의 로그분석기라는 것이 가장 큰 특징이며, 초급 사용자를 위한 배려인 유저별 위험도에 대한 부분을 따로 언급하고 있다는 것이다. 그러나 리눅스의 모든 로그파일을 분석한 것은 아니며 중요 로그만을 분석하였다는 점, 분석되는 서비스에 대한 한계를 가지고 있다. 그러므로 앞으로는 분석 가능한 로그파일과 해당 시스템의 서비스를 늘려 나가야 한다. 또한 보다 편리한 분석과 관리를 위해 더 나은 기능들이 추가되어져야 할 것이다.

[참고문헌]

[1] 신현준 저, "리눅스 서버 구축 & 보안", 사이버출판사, 2002
[2] http://www.superuser.co.kr/linux/logmonitoring/ 추
[3] http://www.acrasoft.com/lw.html
[4] http://lecture.donghae.ac.kr/lab/07-logscanner/log%20scanner.htm
[5] http://kltp.kldp.org/stories.php?story=04/02/06/6420533
[6] 이강성 저, "python", FREELEC, 2003
[7] 최재영, 최종영, 유재우 공저 "프로그래머를 위한 Java2", 홍릉과학출판사, 2003
[8] 강기봉 저, "(시스템 관리자를 위한)리눅스 바이블 9.0", 북스피아, 2003