모바일 웹 서버 로그 분석기 구현†

(Implementation of Mobile Web Server Log Analyzer)

박 홍 진*
(Hong-Jin Park)

요 약 웹 서버가 이상이 발생되면 로그 파일은 에러를 찾는 매우 중요한 정보이다. 웹 서버 분석기는 웹 서버의 로그 정보를 분석함으로써 웹 서비스를 향상시키기 위해 중요한 역할을 수행한다. 그러나 대부분의 기존 웹 서버 분석기는 유선 기반이기 때문에 시공간적 제약성과 이동성에 있어서 문제점이 있다. 본 논문은 기존 유선 중심의 웹 서버 로그 분석기의 문제점을 해결하기 위해 모바일 기반에 웹 서버 분석기를 구현한다. 웹 서버 분석기를 모바일로 구현할 경우, 언제 어디서든지 쉽게 웹 서버 로그정보를 확인할 수 있을 뿐만 아니라 실시간으로 로그파일에 접근하여 최근의 정보를 바로 확인이 가능하다.

핵심주제어 : 모바일, 로그 정보, 웹 서버 분석기

Abstract If web server occurs any fault, log file is important information to find out occurred error. The web server analyzer plays a vital role in improving the web service quality by analyzing log files stores in the web server. But, most conventional web server analysis programs are provided over wired networks, there are problems to be dealt with in terms of time-space restrictions and mobility. In order to solve above problem, this paper represent web server analyser based mobile. In case web server manager based mobile, it can immediately know recent log information accessing real-rime as also anytime, anywhere.

Key Words : Mobile, Log Information, Web Server Analyzer

1. 서 론

최근 정보 기술과 인터넷 기술의 급속한 발전은 컴퓨팅 패러다임 변화는 물론 비즈니스 환경에도 크게 변화를 가져오고 있다.

웹 서버를 기반으로 다양한 서비스를 제공하는 웹 서비스(web services)는 기존의 복잡한 비즈니스 업무를 효율적이고 신속하게 처리하여 경제성, 신속성, 신뢰성, 기능성 등의 다양한 장점을 제공하고 있다. 이러한 웹 이용자에게 웹 서비스를 지속적이고 안정적으로 제공하기 위해서 웹 서버의 실시간적인 모니터링 기술이 필요하다. 웹 서버 관리자에게 중요한 작업 중 하나가 현재의 웹 서버 상태를 파악하여 웹 서버의 안정성을 유지시키는 일이며 또한, 현재 웹 상태 모니터링하고 분석하여 향후의 웹 서버의 안정성 및 확장성을 계획 및 예측할 수 있다[1-7].

웹 서버 관리자는 웹 서버의 트랙픽, 에러(error), 데이터 전송량 등을 실시간적으로 분석하여 보다 효율적이고 안정적으로 웹 서비스를 제공될 수 있도록 해야 한다[8]. 특히, 갑작스러운 사용자 폭증이나 불법적인 사용자가 웹 서버에 접근하면 웹 서버 부하가 짧은 시간 안에 크게

† 이 논문은 2005년도 상지대학교 교내 연구비 지원에 의한 것임.
* 상지대학교 컴퓨터정보공학부

증가하게 되거나 그에 따른 오류를 통해 웹 서버에서 제공되는 서비스가 중단되는 현상이 발생될 수 있기 때문에 웹 서버 관리자는 지속적이고 실시간적인인 웹 서버 모니터링을 수행하여 문제점 발생이나 불안정한 상태로 징후가 발생될시 즉각적으로 상황을 판단을 하여 상황에 맞는 조치가 신속하게 취해야 한다.

웹 서버 시스템에 접속한 이용자들의 행위들을 저장해 놓은 기록 정보가 웹 서버 로그(log) 정보이다[5]. 현재 웹 서버 분석기들은 대부분이 로그에 저장되어 있는 사용자 정보, 방문 시간, 사용중인 웹 브라우저 종류, 방문 페이지, 다운로드 용량, 서버의 에러원인, 8단계의 에러메시지 등의 다양한 정보를 웹 서버 관리자에게 제공하고 있다. 이러한 로그정보를 사용하는 웹 서버 분석기는 웹 서버가 올바르게 동작하는지를 판단할 수 있어야 한다. 이미 오류가 발생하여 그 오류 원인을 알지 못하면 웹 서버 서비스를 복구하는데 오류 원인을 찾기 위해 상당한 시간이 소요될 수 있으며, 이는 중요한 웹 서비스를 제공하는 사이트에서는 매우 중요한 기능이다. 또한, 웹 서버의 사용자가 급증하는 시간대와 가장 많이 방문한 웹 페이지를 알 수 있게 하여 웹 서버의 서비스 향상에 도움을 줄 수 있어야 한다. 가장 많이 방문한 웹 페이지를 알게 되면 사용자 측에서 홍보를 목적으로 할 경우 유용하게 사용할 수 있다.

기존 많은 웹 서버 분석기는 대부분의 유선 중심으로 웹 서버 로그 분석기를 제공하고 있다. 유선 기반이기 때문에 시공간적인 자유로움과 이동성에 한계를 지니고 있다. 본 논문은 모바일 웹 서버 분석기를 구현함으로써 무선으로 웹 서버 로그 정보를 분석하여 언제 어디서든지 쉽게 웹 서버 로그정보를 분석할 수 있는 장점을 제공함에 있다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 연구로 현재 사용되고 있는 웹 서버 로그 분석기를 알아보고, 3장에서는 모바일 기반의 모바일 웹 서버 로그 분석기 구성을 살펴보며, 4장에서는 이를 구현하여 5장에서 결론을 맺는다.

2. 기존 연구

로그파일에 저장되어 있는 각종 접속정보 등은 아파치 웹 서버를 비롯한 대부분 서버에서는 CLF(Common Log Format) 기반으로 되어 있거나 사용자에 의해 지정된 형식으로 운영자에게 다양한 정보를 제공하고는 있지만, 로그에 저장된 내용이 웹 서버의 사용량에 따라서 크기가 엄청나고 바로 그 로그에 저장된 정보를 사용하는 대에는 상당한 어려움이 있다. 웹 서버 로그 분서기는 로그파일에 있는 정보들을 웹 서버운영자들이 손쉽게 보고 정보를 판단할 수 있게 도와주는 도구이다. 즉, 웹 서버 로그 분석기는 로그정보를 여러 가지 형태로 통계를 내어 웹 서버를 운영하는데 좀더 향상된 서비스를 제공하게 하는데 도움을 준다. 다음은 현재 운영되는 웹 서버 로그 분석기이다.

2.1 Webalizer[9]

대부분의 리눅상에서 기본적으로 제공되는 로그 분석기이다. C 언어로 작성되어 있어서 빠른 속도와 편리함으로 웹 서버 분석기로 많이 사용된다. Webalizer은 단순한 사용법과 강력한 기능이 많이 있지만 일반적으로 Webalizer을 이용하면 지속적인 모니터링을 할 수가 없기 때문에 모터링 결과물을 웹으로 출력하기 위해서는 Webalizer을 반복적으로 실행해야 하는 단점이 있다.

2.2 Awstats[10][11]

Webalizer와 함께 많이 쓰는 웹 로그 분석기이다. Perl 언어로 구현된 Awstats는 Weablizer와 같게 리눅스 환경에서 수행되는 무료 소프트웨어이고 보다 향상된 인터페이스를 제공하지만 이 로그분석기 역시 실시간적으로 로그파일에 접근하지 않고 일정한 시간마다 정보를 갱신한다. Awstats 분석기는 모바일기반이 아닌 유선기반이고 일정한 시간마다 분석을 하기 때문에 보는 즉각적인 정보를 알 수가 없다는 단점이 있다.

2.3 WiseLog[12]

기능면에서는 Webalizer과 Awstats보다 뛰어나지만 저장되는 로그파일을 하루에 한번통계를 내

기 때문에 웹 서버에 문제가 발생하였을 경우 상당한 시간이 지나야 알거나 웹 서버에 이상이 발생시 알게 될 수 있다는 단점이 있다. 유료 프로그램이기 때문에 개인용도의 웹 서버를 사용하는 사용자는 사용하는데 비용적인 부담을 가질 수 있다.

2.4 Logger[13]

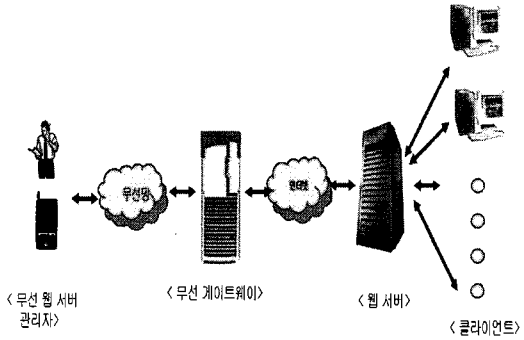
소프트웨어를 따로 설치가 필요 없는 ASP서비스로서 실시간으로 웹 로그 분석기이다. 기존의 것과는 달리 스크립트를 분석하고자 하는 웹 페이지에 삽입하는 것만으로도 분석이 가능하다. 스크립트이기 때문에 로그 분석에 서버 자원이 거의 사용되지 않기 때문에 상당히 효율적인 웹 로그 프로그램이다. 또한, 웹 서버에 생성되는 로그정보를 이용하지 않고 분석하고자하는 웹 사이트의 웹 페이지내의 '로거'의 로그 분석 스크립트(page embedding)를 삽입하는 것으로 로그분석을 위한 데이터를 수집하여 분석한다. 하지만 서비스는 회사측에 있는 서버에서 분석을 하는 것이므로 반드시 비용을 지불하고 서비스에 가입해야한다. 회사측에서 사용자의 로그정보를 볼 수 있기 때문에 중요할 수 있는데 정보가 새어나갈 확률이 높다,

2.5 Analog

윈도우, 맥, 리눅스, 유닉스 그리고 BSD 환경까지 다양한 운영체제를 지원하고 있으며, 대부분의 많은 웹 로그파일 포맷을 인식할 수 있다. 분석된 결과는 하나의 HTML파일로 만들어진다. 하지만 분석된 정보를 다른 로그분석도구와는 달리 이해하는데 어려움이 따르며 기존의 로그분석도구처럼 실시간적으로 로그의 정보를 가져올 수 없다.

3. 모바일 웹 서버 분석기 구성도

기존의 웹 서버 분석기들은 2장에서 설명된 것처럼 장단점을 가지고 있으나 대부분의 유선 중심의 웹 서버 정보를 분석함으로써 이동성과 시간성에 제약성이 지니고 있다. 본 논문에서는 웹 서버 정보를 모바일 환경에서 제공함으로써 웹 서버 관리자가 시공간에 관계없이 웹 서버의 정보를 파악할 수 있는 잇점이 있다. 또한, 로그파일에 실시간으로 접근하여 원하는 정보만을 추출해서 분석을 하기 때문에 기존의 분석도구보다 보다 가장 최근의 로그정보를 확인할 수 있다.



<그림 1> 모바일 기반 로그 분석기 구성도

<그림 1>은 본 논문에서 제시하고 있는 모바일 기반의 웹 서버 로그 분석기의 전체 구성도이다. <그림 1>에서 나타난 것처럼 클라이언트인 웹 서버 이용자들이 웹 서버로 접속을 하면 접속한 이후에는 웹 서버의 로그 정보에 저장이 되고 이러한 정보는 인터넷 프로토콜을 무선 프로토콜로 변경하는 무선 게이트웨이를 통해 무선망으로 정보가 보내지며 최종적으로 무선 단말기를 휴대한 웹 서버 관리자에게 웹 서버 정보가 전송이 된다. 또한 웹 서버 관리자는 무선 단말기를 통해 웹 서버 관리에 필요한 명령어를 전송하여 결과값을 얻을 수 있다.

무선 단말기에 전송되는 정보는 크게 접속 로그정보와 에러 로그 정보로 구분된다. 접속로그 정보는 웹 서버로부터 전송되는 정보로 방문한 사용자가 어떤 링크를 통하여 접속했는지, 일정 기간동안 접속량, 전송 데이터의 양등을 파악할 수 있으면, 특히 웹 서버 접속자들의 많이 접속한 페이지도 파악할 수 있다. 에러 로그 정보는 크게 8단계로 구분되어 지며, 이는 emerg, alert, crit, error, warn, notice, info, debug이다. 최상위의 emerg와 alert, crit 같은 경우는 웹 서버에 이상이 생겨서 서비스가 중단되거나 중단될 수 있는 상태에 있음을 의미한다. 나머지 단계는 서버에 이상을 주지는 않지만 조기에 조치를 하면 심각

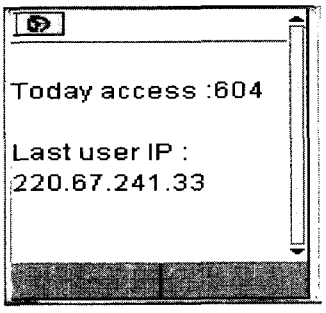
한 에러가 발생하는 것을 예방할 수도 있다. 이러한 오류 정보를 통해 웹 서버 관리자는 외부로부터의 잘못된 웹 서버 접속에 대한 정보를 분석함으로써 보다 안전한 웹 서버 유지를 유지시킬 수 있다. 이러한 각 에러의 단계와 의미는 <표 1>과 같다.

<표 1> 웹 서버의 8단계 에러 메시지

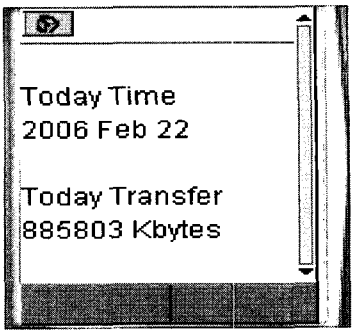
에러 단계	에러의 의미
Emerg	불안정한 시스템 상황
Alert	즉각적인 조치 필요
Crit	중대한 에러
Error	비교적 중대하지 않은 에러
Warn	경고
Notice	중대한 것은 아닌 일반적인 메시지
Info	정보
Debug	디버그 레벨

4. 모바일 웹 서버 분석기 구현

본 논문에서 구현한 웹 서버 로그 분석기는 크게 웹 서버 접속 정보를 파악할 수 있는 접근 로그 정보와과 웹 서버 접속 오류 정보를 파악하는 오류 정보 추출로 구성되어 있다. 현재 대부분의 유선 기반 웹 서버 로그 분석기는 웹 서버 접속 정보를 위주로 추출하지만 본 논문에서는 향후 웹 서버의 잘못된 사용에 대한 분선을 위해 오류 정보도 추출한다.

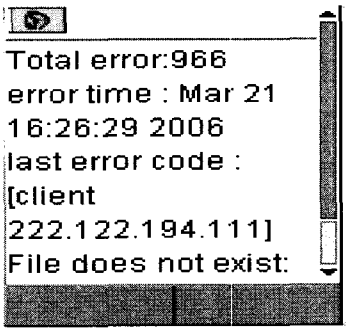


<그림 2> 당일 웹 서버 접속 총 횟수와 마지막 접속 IP 주소

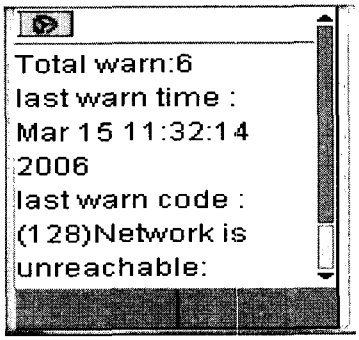


<그림 3> 당일 데이터 전송량

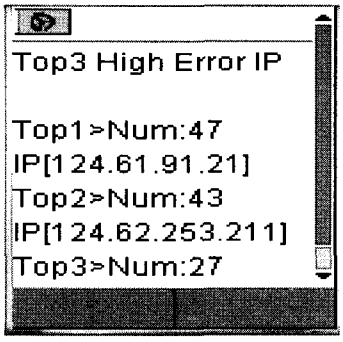
<그림 2>은 당일날짜로 접속한 사용자 수와 마지막으로 접속을 한 접속자의 IP를 보여주는 화면이다. 이같이 하루단위의 접속량을 통계적으로 사용하면 월단위의 접속량을 알 수가 있다. 이는 통계적으로 한달 중 가장 많이 접속량이 몰리는 날을 미리 예측할 수 있는 장점이 있다. <그림 3>는 당일의 현재까지의 전송량을 나타낸 것이다.



<그림 4> 마지막 에러 정보



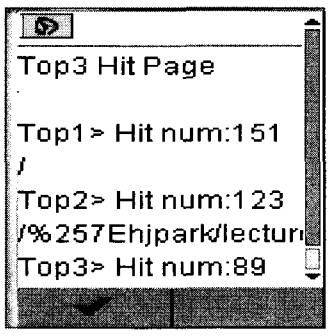
<그림 5> 마지막 경고



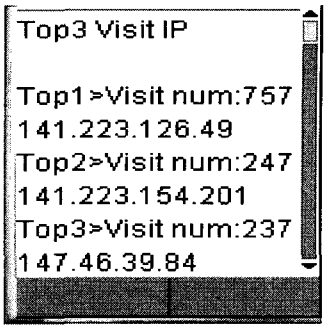
<그림 6> 에러 발생 IP 상위 3개

<그림 4>은 에러 8 레벨에서 error_level을 나타낸 것으로 당일 error_log 발생량과 마지막 발생시간 또, 마지막 error_level의 에러코드를 나타낸 것이다. <그림 5> 에러 8 레벨에서 경고(warm_level) 정보를 타나낸 것으로 error_level과 같이 경고의 발생량, 마지막 발생기간, 마지막 에러코드를 나타낸다. <그림 6> 에러 8 레벨에서 가장 많이 에러를 발생시키는 상위 IP 주소를 나타낸다.

<그림 7>은 웹 서버 파일 중에 가장 많이 히트(hit)한 페이지를 추출해서 보여준다. 이 정보를 이용하면 사용자가 어느 페이지를 가장 많이 방문하였는지를 알 수 있다. 그 점을 서비스에 응용하면 이벤트 등을 할 경우 배너 광고를 가장 많이 방문한 페이지에 노출시켜 이벤트 참여율을 높일 수도 있고, 배너 광고를 할 경우 이 정보를 토대로 할 수 있는 장점이 있다.



<그림 7> 당일 상위 히트 페이지



<그림 8> 가장 많이 웹 서버 IP

<그림 8>은 웹 서버에 가장 많이 방문한 IP 정보를 나타낸다. 가장 많이 접속한 정보로 어떠한 사용자가 가장 활발히 활동을 하는지를 알 수가 있다.

5. 결 론

기존의 웹 서버 로그 분석기는 유선 중심이기 때문에 이동성에는 제약을 받으며 웹 서버의 상태를 지속적으로 확인해야 한다. 또한 로그파일에 실시간적으로 접근하는 방식보다는 일정시간단위로 접근하여 통계를 낸다. 이는 로그파일의 크기가 너무 크다는 점 때문에 분석시에 많은 시스템 자원을 사용하기 때문이다.

하지만 본 논문에서는 일정시간의 로그파일만을 실시간적으로 접근하여 분석할 수 있으며, 언제 어디서든지 실시간적으로 로그파일의 정보를 확인이 가능하다. 로그파일을 일정시간 간격만을 분석하는 것은 아주 장시간 서버를 관리하지 않는 일이 발생하지 않기 때문이다. 이러한 점에서 모바일기반의 로그분석기는 실시간적으로 최근의 서버의 로그파일에서 문제점을 알 수가 있다. 이는 기존의 로그분석기보다 서버에 심각한 에러가 발생시 빨리 발견하여 서버를 정상화 시킨다. 이러한 이동성과 실시간적인 점이 기존의 웹 서버 분석기보다 장점이 된다.

참 고 문 헌

[1] "Web Server Monitoring", White Paper, http://www.freshtech.com/WhitePaper.htm

[2] Web Site Monitoring, http://www.digitalventure.net/vp/index.html
[3] Xiaozhe Wang, Ajith Abraham and Kate A. Smith, "Intelligent web traffic mining and analysis", *Journal of Network and Computer Applications,* Volume 28, Issue 2, April 2005.
[4] Martin M. Echols, David K. Smith and David S. Nirschl, "Web-based instrument monitoring system", *Journal of the Association for Laboratory Automation,* Volume 9, Issue 6, December 2004.
[5] E. Anderson, D. Patterson, "Extensible, Scalable Monitoring for Clusters of Computerts", *Proceedings of 1997 LISA Conference,* 1997.
[6] A Perl Script for Monitoring Apache Server Status, http://webreview.com/pub/1999/02/webm/index.html
[7] MARS(Monitoring Application for Resource and Server),http://www.altara.org/mars.html
[8] http://www.apache.org
[9] http://www.mrnuix.net/webalizer
[10] http://awstats.sourceforge.net
[11] http://www.awstats.org
[12] http://www.wiselog.com
[13] http://www.logger.co.kr

박 홍 진 (Hong-Jin Park)

- 정회원
- 1993년 2월 : 원광대학교 컴퓨터공학과 (공학사)
- 1995년 8월 : 중앙대학교 컴퓨터공학과 (공학석사)
- 2001년 8월 : 중앙대학교 컴퓨터공학과 (공학박사)
- 2001년 9월 ~ 현재 : 상지대학교 컴퓨터정보공학부 조교수
- 관심분야 : 운영체제, 분산시스템, 모바일