



로그

분석의 이해와 방법(2)

1	로그 분석 도구의 종류	1
2	WebTrends	1
3	Analog	3
4	WebAlizer	4
5	Weblog	5
6	FastStats	7
7	로그 분석기 직접 만들기	7
8	로그 분석 벤치마크	8
9	로그 분석 실습	9
10	요점 정리	10

1. 로그 분석 도구의 종류

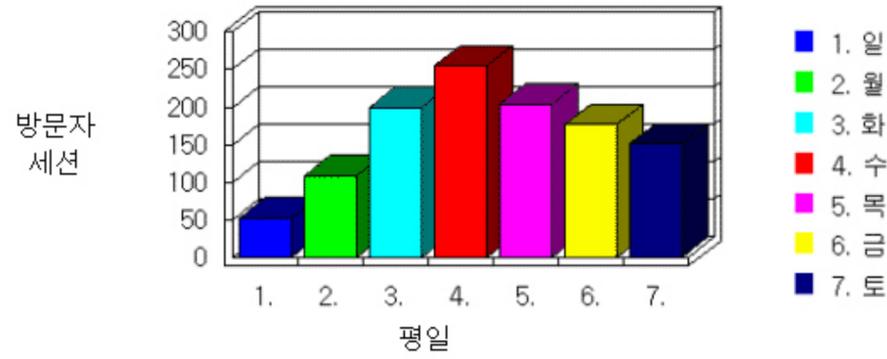
3Dstats	IStats
Access Panda	JV's WWWStat
AccessWatch	MachHTTP Utilities
Analog	Metasummary
BrowserCounter	MK-Stats
Cardiff's Daily Graphs	mw3s
Combined Log File System	PressView
CreateStats 1.0	Raytraced Access Stats
EIT's Getstats	RefStats 1.0
Free Aaddzz Counters and Trackers	Suspended Animation Club Statbot
FTPWebLog	UCI wwwstat
getgraph.pl	Utilities
getstats_plot	WAMP
Graphic wwwstat - gwstat	WebLog
httpd_log 2.0	WebStat
iisstat	WebTrends
	wusage 2.4

2. Web Trends

1 이해

- 홈페이지
- WebTrends(www.webtrends.co.kr/default.htm)
- Sample(온라인 메뉴얼)
www.webtrends.co.kr/Sample/Log-Analyzer/report.htm
- 다양한 웹 서버에 대해 원격 또는 로컬로 로그 분석 가능함
- 내장된 원격 보고용 서버를 이용하여 자동적으로 고객들의 데스크 톱 웹브라우저를 이용하여 접근할 수 있는 중요한 보고서를 자동적으로 생성함
- 웹 서버에서의 트래픽 측정, 웹 사이트 방문자의 숫자 및 수준 분석, 웹 사이트 방문자의 브라우징 경로 및 웹 사이트에 관한 전반적이 정보 분석에 탁월함
- 종류별로 패키징되어 있으며 세어웨어를 다운 받을 수 있음

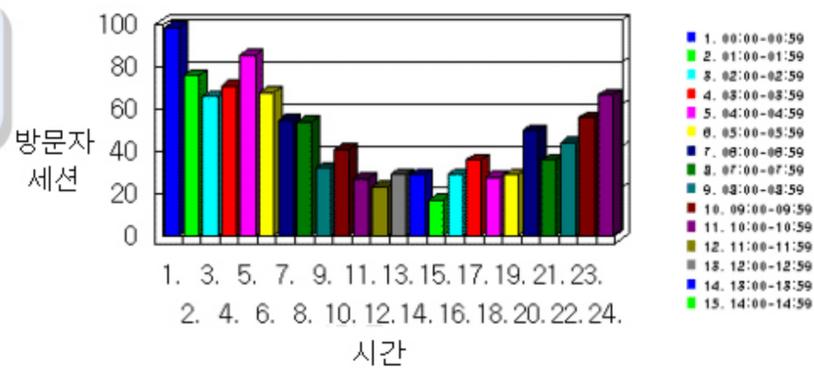
2 주 중 날짜별 활동 레벨



날짜	히트	히트 합계에 대한 점유율(%)	방문자 세션
1 일	423	4.2%	52
2 월	1,094	10.86%	109
3 화	1,908	18.94%	200
4 수	1,947	19.33%	255
5 목	1,850	18.37%	204
6 금	1,557	15.46%	177
7 토	1,290	12.81%	151
평일 합계	8,356	82.98%	945
주말 합계	1,713	17.01%	203

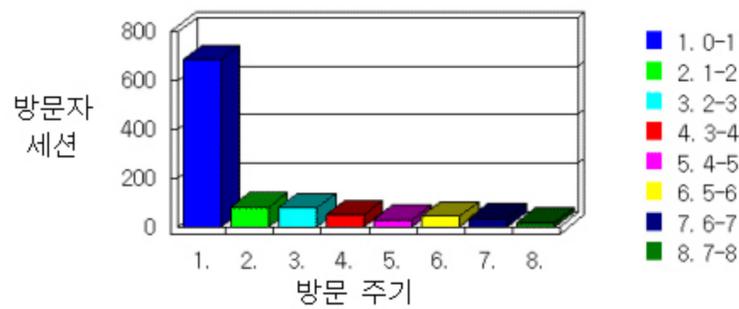
3 시간 데이터 의미

시스템 정비를 위해 가장 좋은 시간이 몇 시인지를 결정하는데 유용함



4 방문 길이별 활동 레벨

방문자들이 얼마나 오랫동안 웹 사이트를 보았는지를 측정하는데 유용함



3. Analog

1 이해

- 홈페이지
Analog(www.statslab.cam.ac.uk/~sret1/analog/)
- Sample(온라인 메뉴얼)
www.statslab.cam.ac.uk/~sret1/stats/stats.html
- 무료이며 로그 사이즈가 커도 빠른 시간 안에 처리한다는 것이 특징
- 35가지의 언어로, 여러 가지의 리포트가 제공됨(Korean kr.Ing)
-There are 35 languages included in version 4.16



Analog

The most popular logfile analyser in the world

Introduction
Ultra-fast
Scalable
Highly configurable
Reports in 34 languages
Works on any operating system
Free software

2 General Summary

(Go To: Top: General Summary: Monthly Report: Daily Summary: Hourly Summary: Domain Report: Organisation Report: Search Word Report: Operating System Report: Status Code Report: File Size Report: File Type Report: Directory Report: Request Report)

Monthly Report

Each unit (-) represents 20,000 requests for pages or part thereof

Jan 2000:	710,600:	293,076:	13.670:	
Feb 2000:	716,730:	308,201:	14.410:	
Mar 2000:	737,935:	322,999:	13.712:	
Apr 2000:	634,589:	281,790:	11.813:	
May 2000:	710,518:	312,266:	13.140:	
Jun 2000:	638,304:	267,655:	13.046:	
Jul 2000:	758,029:	337,284:	12.784:	
Aug 2000:	863,333:	409,852:	12.402:	
Sep 2000:	911,895:	404,312:	14.951:	
Oct 2000:	1,064,720:	460,286:	20.295:	
Nov 2000:	1,058,820:	463,057:	18.890:	
Dec 2000:	869,452:	335,833:	15.924:	
Jan 2001:	1,066,684:	399,606:	19.899:	
Feb 2001:	1,068,341:	397,145:	20.011:	

3 Daily Summary

Each unit (-) represents 50,000 requests for pages or part thereof

day:	pages:	
---	-----	
Sun:	1,542,334:	
Mon:	1,846,979:	
Tue:	2,016,134:	
Wed:	1,991,786:	
Thu:	1,974,427:	
Fri:	1,844,896:	
Sat:	1,513,197:	

4. WebAlizer

1 이해

- 홈페이지
Webalizer (www.mrunix.net/webalizer)
- Sample (온라인 매뉴얼)
www.mrunix.net/webalizer/sample/index.html
- Unix, Linux, Solaris, Macintosh, Windows 95/98/NT 등의 모든 운영체제의 Apache, Microsoft IIS, WebStar와 같은 모든 타입의 웹 서버에서 사용할 수 있는 로그 분석기
- 트래픽을 3D bar로 나타내며, 방문과 추적에 대한 기록, 레퍼링 사이트를 분석하는 기능 등의 많은 기능을 제공하는 로그 분석기

The Webalizer
What is your web server doing today?



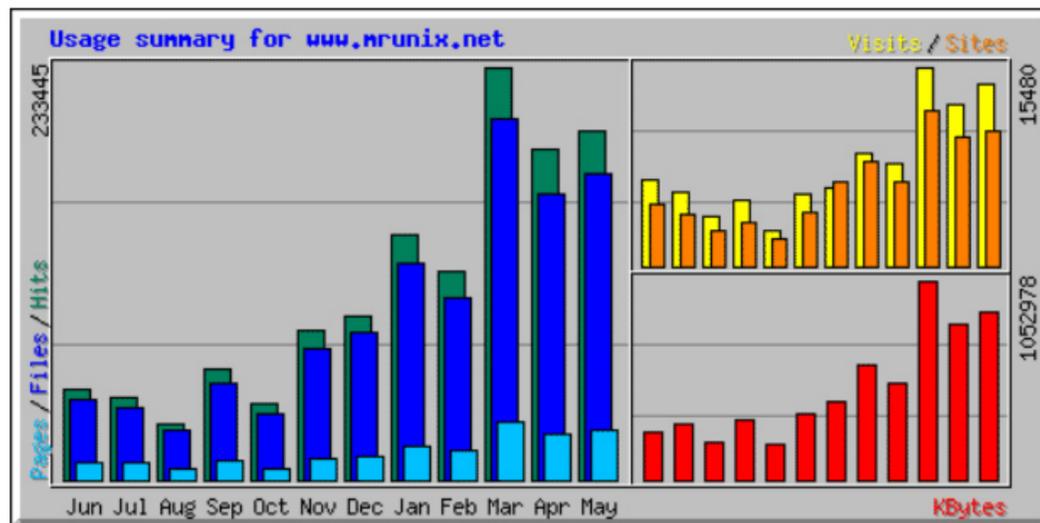
2 통계 및 분석

Usage Statistics for www.mrunix.net

Summary Period: Last 12 Months

Generated 23-Oct-2000 02:17 EDT

[Daily Statistics] [Hourly Statistics] [URLs] [Entry] [Exit] [Sites] [Referrers] [Search] [Agents] [Countries]



5. Weblog

1 이해

- 홈페이지
www.naro.co.kr/software/weblog/index.htm
www.weblog.com/kr/index.html
- Sample (온라인 매뉴얼)
www.naro.co.kr/software/weblog/guide/guide.html

서버 사용 요건

서버 환경	최소 사양	권장 사양
윈도 NT	P166MHz / 64MB RAM / 윈도 NT4.0	PIII 500MHz / 128MB
리눅스	P166MHz / 64MB RAM	PIII 500MHz / 128MB
SUN	SPARCSTATION 10 / 64MB RAM	SOLARIS 2.6
HP	HP715 / 64MB RAM	HP-UX 10.0
IBM	IBM RS6000	IBM AIX
DEC	ALPHASERVER1200 / 128MB RAM	DEC OSF1 V4.0
FreeBSD	P166MHz / 64MB RAM	FreeBSD 2.2.5

2 특징 분석

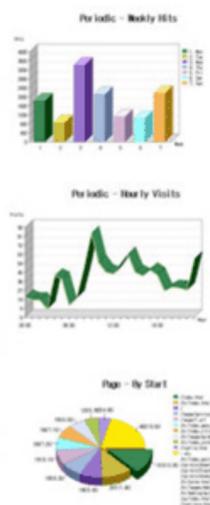
웹 로그는 국내에서 자체 개발된 통계 분석용 프로그램이다. 특히 접속자수의 정확한 산출, 방문객의 거주지를 지도와 함께 그래픽으로 표시, 다양한 그래픽 보고서 작성, 한글 통계 용어의 사용 등 다양한 점들을 보유하고 있다. 웹 서버가 동작하는 거의 모든 OS를 지원한다. 데모를 무료로 사용해볼수 있다.

<제품의 특징>

- 접속수와 실제 방문객 수의 구별
- 방문객의 거주지를 시도별로 분류
- 방문객의 접속 ISP별 집계
- 홈페이지 디렉터리와 파일별 통계
- 홈페이지 메뉴 분석
- 기간별(월, 주, 요일, 일, 시간) 그래프
- 접속 시간과 사용 데이터 양으로 요금 계산
- 전세계 국가별/업종별 통계
- NT, 리눅스, 유닉스 지원

3 보고서의 종류

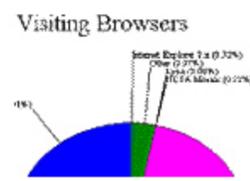
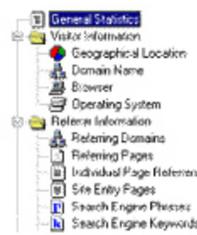
기본 보고서	에러 분석	방문 횟수	방문 현황 G
일반 현황 분석	방문 시간대별 분석	이동 데이터 방문 현황 G	이동 데이터 방문 현황 G
월 사용 현황 G	요일별 방문 현황 G	최근 접속자 방문 현황 G	최근 접속자 방문 현황 G
월 접속 현황 G	시간대별 방문 현황 G	가나다순 ISP 목록	
월 방문 현황 G	방문생성 횟수 방문 현황 G	페이지 분석	
월 에러 현황 G	방문생성 횟수 방문 현황 G	전 접속 횟수 방문 현황 G	
월 데이터량 현황 G	방문생성 횟수 방문 현황 G	소요 시간 방문 현황 G	
요일별 사용 현황 G	방문생성 횟수 방문 현황 G	최근 접속자 방문 현황 G	
요일별 접속 현황 G	방문생성 횟수 방문 현황 G	방문시작 횟수 방문 현황 G	
요일별 방문 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
요일별 에러 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
요일별 데이터량 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
평일/주말 비교표	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
평일/주말 비교표 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
일일별 사용 현황	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
일일별 접속 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
일일별 방문 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
일일별 에러 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
일일별 데이터량 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
시간대별 사용 현황	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
시간대별 접속 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
시간대별 방문 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
시간대별 에러 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
시간대별 데이터량 현황 G	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
시스템 분석	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	
소개 사이트 목록	방문생성 횟수 방문 현황 G	방문내거 횟수 방문 현황 G	



6. FastStats

1 이해

- 1분에 700MB 의 로그를 처리해서 분석하는 로그 분석기이다.
FTP나 Web으로 로그 파일에 접근할 수 있는 컴퓨터라면, 어디에서나 실행 가능하다.
아파치를 포함, NCSA httpd, CERN httpd, 마이크로소프트 IIS, 오렐리의 Website와 매킨토시 서버 등 다양한 웹 서버를 지원한다.
- Mach5 (mach5.com)
- Sample (온라인 메뉴얼)
mach5.com/products/faststats/report.html



2 FastStats vs WebTrends

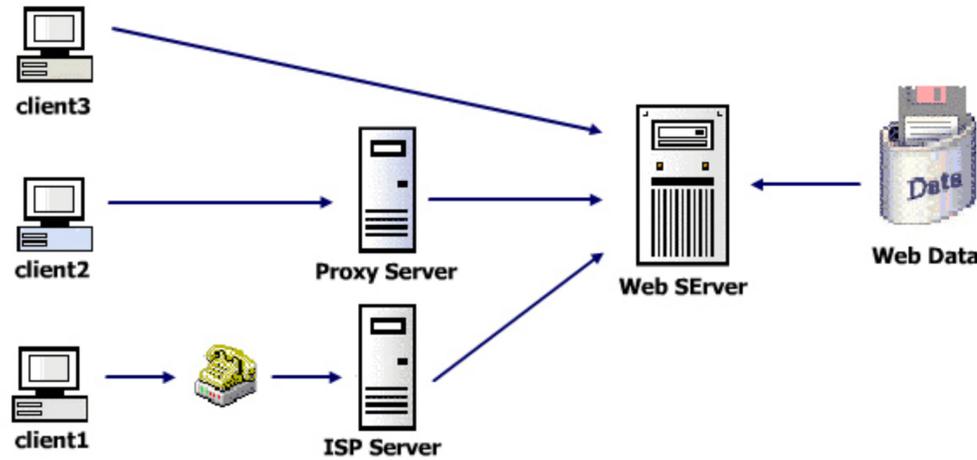
Feature	FastStats Analyzer	WebTrends Log Analyzer
WebTrends Log Analyzer	\$99.95	\$499
Speed	100-200 MB per minute (up to 700mb/min on some computers)	10-30 MB per minute (this is the advertised speed, it's much less in real life)
Hyperlink Tree View	Yes	No
Individual Page Referrer Report	Yes	No
Interactivity	Yes	No
File Tracking Reports	Yes	No
Multi-threaded Reverse DNS	Yes	Yes (Kindof)

7. 로그 분석기 직접 만들기

- Why?
- 기존 로그 분석기로는 분석하고자 하는 내용이 불충분할 경우가 많다.
- 로그파일의 구조만 분석하면 비교적 간단하다.
- 내가 필요한 정보로 만든다.

웹 로그는 정확하지 못하다?

- 프록시 서버의 IP Address만을 기록
- 배너나 웹 문서가 캐시에 보관되어 있는 경우
- 브라우저의 이미지 보기 기능을 오프시킨 경우
- 페이지가 로딩되기 전에 취소된 경우



8. 로그 분석 벤치마크

	로그 사이즈	분석 시간	CPU 점유율	환경 설정의 용이성	사용자 환경	분석 결과 에 대한 설 명	그래픽 지원	고급 로그 분 석 기능	NT / Unix
Web Trends	254MB	10분	95%	쉬움	GUI	자세함	지원함	가능	NT Only
Analog	254MB	1분 30초	80%	어려움	Text	없음	지원함 (일부)	가능	NT / Unix
Web alizer	254MB	1분 40초	85%	어려움	Text	없음	지원함 (일부)	가능	NT / Unix
Web Log	254MB		97%	쉬움	Text	자세함	지원함	가능	NT / Unix
WWW Stat	254MB	17분	95%	어려움	Text	없음	지원 하 지 않음	가능	Unix

	템플릿 제공	한글 지원	고급 분석 기능	DB 지원	배너 분석 기능	상용 여부	기술 지원 여부
Web Trends	제공	지원	지원	지원(ODBC)	지원	상용	지원
Analog	없음	지원 안함	지원 안함	지원 안함	지원 안함	무료	불가
Web alizer	없음	지원 안함	지원 안함	지원 안함	지원 안함	무료	불가
Web Log	제공	지원	지원 안함	지원 안함	지원 안함	상용	지원
WWW Stat	없음	지원 안함	지원 안함	지원 안함	지원 안함	무료	불가

평가	분석 속도	편리성	분석 결과 만족도	전체적인 순위
Web Trends	4	1	1	1
Analog	1	3	3	3
Web alizer	2	4	4	4
Web Log	5	2	2	2
WWW Stat	3	5	5	5

<선호도 조사>

지난 2000년 5월 1일부터 5월 17일까지 17일간 웹매니아 회원들을 상대로 조사한 로그 분석 툴에 대한 선호도 조사의 결과는 다음과 같다.



■ 테스트 환경

NT : 펜티엄III 500MHz / 128MB 메모리 / NT 4.0 / 서비스 팩 5 / 노트북

리눅스 : 펜티엄II 300MHz / 128MB 메모리 / 레드햇 리눅스 6.0

* DNS LookUp은 모두 안하는 것으로 테스트

■ 평가

전체적으로 웹로그와 웹트렌즈가 분석 결과 및 사용자 편의성에서 신경을 많이 쓴 부분이 돋보인다. 웹트렌즈의 경우 보고서 유형별 템플릿을 지원해 사용자로 하여금 쉽게 보고서를 작성할 수 있게끔 하고 있으며 분석 속도면에서는 Analog와 WebAlizer 두 개의 제품이 가장 속도가 빨랐다. 여러 제품이 분석 결과에 그래픽을 지원해 사용자로 하여금 한 눈에 쉽게 알아 볼 수 있도록 하였다. 분석 속도는 조금 느렸지만 분석 결과 및 고급 분석 기능 / 사용자 편의성이 뛰어난 웹트렌즈를 가장 추천한다. 특히, 단순 웹 로그 분석 뿐 아니라, 서버 모니터링도 같은 고급 정보를 제공하는 웹트렌즈 엔터프라이즈 수트를 사용하면 더 많은 정보를 얻을 수 있다.

■ WebTrends를 다운 받아 로그 분석을 실행한다.

- 1) 하루 평균 페이지 뷰/히트 수/유일 방문자 수?
- 2) 방문객들이 어느 사이트를 통하여 가장 많이 접근하며, 얼마나 많은 사용자들이 방문하였는가?
- 3) 메인 페이지 이외에, 방문객들이 가장 많이 접근하는 경로는?
- 4) 이 사이트를 찾을 때 가장 많이 사용되는 검색엔진의 키워드는?
- 5) 방문자들은 하루 중 어느 시간대에 가장 많이 방문하는가?
- 6) 방문자들은 주 중 어느 요일에 가장 활발한 활동을 보이는가?
- 7) 방문객들의 브라우저 사용 분포는?

10. 요점 정리

- 웹 로그를 분석하기 위한 다양한 소프트웨어들이 나와 있다. 기능 차이도 있고 또 가격은 무료부터 수천만원 대까지 다양하다. 단순한 트래픽 분석이라면 무료를 이용해도 충분하다. 하지만 좀 더 세밀한 분석을 원한다면 고가의 제품을 구입해야 한다.

결과는 참조만 하고 절대적으로 믿으면 안된다. 좀더 정확한 결과를 얻고자 한다면 여러 번 필터링 해야 할 것이다.

- 국내 통계 사이트

- KNP(Korea Netizen Profile) knp.adic.co.kr/uw/dispatcher/knp/index.html
- 인터넷 메트릭스 www.internetmetrix.com
- 한국 인터넷 정보 센터 (www.nic.or.kr)
- 인터넷 마케팅통계청 (www.marketcast.co.kr)

- 국외 통계 사이트

- eStats(www.emarketer.com/welcome.html)
- Forrester Research(www.forrester.com)
- International Data Corporation (www.idc.com)
- Jupiter Communications(www.jup.com/home.jsp)
- Nielsen//NetRatings(www.nielsen-netratings.com)
- CyberAtlas (cyberatlas.internet.com)