

[Home](#) >> [CyberStrategy](#) >>

Web Log Analyzers

Since the website started up in 2000, GlobalSecurity.org has been using [HTTP-Analyze](#) to translate the server access logs into tables that can be read and understood by humans. However, our traffic has grown so much that the application can no longer swallow the large size of our log files. You can see that we've had erratic [traffic reporting](#) since May 2003.

Server logs contain valuable information about site visitors and your site's contents: numbers of visitors, dates, times, types of browser used, which sites and search engines are referring visitors to your site, what words or phrases visitors have used in search queries, pages viewed, and other elements of usage statistics. When condensed into periodic reports, this information can assist you in making decisions about your site: Do you need to upgrade your server? Do you have broken links? Should you stop doing design tricks for the old Netscape Navigator 3 browser? In addition, if you are serving ads on webpages, it can assist you because the advertisers want to know what type of visitor you have. For instance, look at this ZDNet article, [Web traffic analysis could save your site](#). In brief, web analysis tools can be invaluable in appraising the effectiveness of a website and track and in defining and targeting your audience.

There are many alternatives in the market, depending on the operating system, feature set, sophisticated reporting requirements and additional functionality. Many of the high-end products are for e-commerce and visitor tracking and may require additional modules for these features. Web analytics has expanded to a broad range of services, like load balancing across a cluster of servers and calculating conversion rates (how frequently visitor becomes a purchaser of goods or services) and return on investment.

There are four ways of analyzing your traffic:

- **Server processing:** the application is installed on the web server and periodically the access log file is processed and a report is created, usually a web page. The selection of the appropriate program for your needs will also depend on how comfortable you or your staff may be with installing and configuring an application on a web server remotely. Sometimes, the program

Resources

- [Counterpane's Log Analysis Resources](#) - more stuff than you can get through in a week, closely related to Internet security. Log files are found in all servers and network devices, not just web servers, and provide a valuable auditing tool for security.
- WebMonkey: [Tracking Your Visitors](#) ::: [Logfile Lowdown](#) ::: [Who's Linking to You](#)
- [Traffic Analysis Primer](#)
- [Patrick's Web Performance Tuning](#)
- Andrew B. King's [Web Optimization](#)
- Rex Baldazo - [Managing and parsing your Apache logs](#) Builder.com, 2003
- Jean-Luc David - [Gather Web server stats with a custom PHP app](#) Builder.com 2003
- Chris Beasley - [It's a Hit! Gauging Success through Traffic Analysis](#) Sitepoint.com
- Charlie Morris - [There's gold in them there log files!](#) (WDVL)
- Cliff Wootton - [Analyzing Log Files](#) Web Developer's Journal
- Bruce Morris [Software for Analyzing Your Web Site Traffic](#) Web Developer's Journal, 1999
- Apache Guide: Logging -- a tutorial by Rich Bowen. [Part 1: Understanding Your access log](#), [Part 2: Error Logs](#), [Part 3: Custom Logs](#) [Part 4: Log-File Analysis](#) and [Part 5: Advanced Logging Techniques and Tips](#)
- Rob Tougher - [Apache Log Analysis Using Python](#), Linux Gazette, November 5, 2002
- **Web Directories:** [Yahoo](#) ::: [Open Directory Project](#) ::: [Business.com](#) ::: [HotScripts.com](#) ::: [LookSmart](#)

provided by your hosting company is sufficient for your needs, especially if you are just getting started.

- **Desktop processing:** the program is installed on a PC. The access log file is downloaded from the web server on a daily, weekly or monthly basis, and the program creates the report.
- **Outsourced:** traffic analysis is done by a service provider who usually inserts a pieces of JavaScript code into each page. At the low end, this may be a simple hit counter system, that links to a centrally hosted web page with information about visitors. At the other end of the spectrum, services can produce reports as web pages or e-mailed as MS Excel spreadsheets or PDF files.
- **Independent tracking:** you can also use services like Nielsen/NetRatings, or Media Metrix to provide independent figures. These companies actually follow users around or monitor traffic through specific network pipelines.

All methods have their pros and cons. When working with large log files, you don't want to have to download a gigabyte file so a server solution saves the pain by generating the report. However, server applications require special configuration. This exercise is not for the faint of heart. Alternatively, your web hosting service may be able to provide support, though it may come with a price if the web log analyzer is not standard on your service.

All of the commercial products come with a set of standard reports, but they also allow the user to customize a report that meets specific needs.

Pricing can also vary according to the business model of the provider and the use of the application. For instance, a hosting company may buy hundreds of licenses for its customers' websites, and roll the cost into the monthly bill. Some products are priced according to the traffic volume because a more robust application is needed to handle gigabyte-sized log files. This is the case of [IndexTools](#) -- a million pageviews can result in a charge of \$500 a month. [WebTrendLive](#) is another service of this kind. In the list below, we have focused on products that can be implemented on a single server.

Open Source or Freeware

Most of these products require a Net-savvy administrator who can install the programs and configure them either from the commandline or with text configuration files. Because many of these programs have been around for a while and installed on thousands of servers, you can usually find assistance (free or low-cost) when you run into trouble.

- [Analog](#) is one of the most common log analyzers on the Web. See this brief [PCMag article](#). [ReportMagic](#), also free, can help you spruce up the format.
- [Funnel Analyzer](#) is reviewed briefly in [PCMag](#) (Linux, Solaris). Also check out the [Web Profiler](#) application for site mapping and evaluation, which is currently available free of charge (See [this review](#) for more info). These applications are actually far more sophisticated than other freeware products. The company, Quest Software, has moved on to other areas, like database, application performance and infrastructure management, as well as a high-end analysis programs. The web traffic products are used a promotional tools for the more advanced applications. See a Web Analyzer [sample report](#). There is an excellent [support site](#) that provides additional information on log analysis, as well as a knowledge base and an online forum. This is a real steal since they used to sell for hundreds of dollars.
- [Webalizer](#) - unpolished graphics and format, looks a lot like [HTTP-Analyze](#). Very geeky install process. Only works on Linux or Unix, but it is used by many ISPs and web hosting services.

- [AWStats](#) - See this Builder.com article, [Handle log file analysis with AWStats](#). Here is a [comparison chart](#) of AWStats, Analog, Webalizer and Hitbox. Check out the [demo page](#).

Low Cost

- [Sawmill](#) \$99 A [sample report](#). This has a strong following among moderate-size websites.
- [WUsage](#) \$25 for single site license. It's pretty bare-bones, even clunky.
- [WebLog Expert](#) - a lite version is free for non-commercial use, and the full version is \$75. The program resides on a PC and downloads logs from the server.
- [LogJam](#) is a Windows application. \$150. Requires some familiarity with the MySQL database, but allows you to customize all the reports. It maxes out at about 50,000 pageviews per log. It requires a Windows web server. The company slogan is "Web Traffic Analysis for the Anal Retentive in All of Us." Be forewarned.

Commercial

What do you get by laying down hard cash (or credit card) for an application? Consistent support, a product that has been tested in a range of environments and stands a chance of being around for a while, and the right to complain when the product does not work as advertised.

- [Urchin](#) \$895 for the Base Module. Sites using this application include [NASA](#), [Verio/NTT](#) and [Fast Company](#). See the [demo page](#).
- SurfStats [Log Analyzer](#) - product line ranges from \$95 for the Standard Edition to \$795 for the Enterprise Edition. See a [demo page](#).
- Sane Solutions [NetTracker](#) \$495 for the Professional version and \$995 for the Enterprise edition, though their website seems to indicate that the Enterprise edition could not handle the traffic of this site. There is an eBusiness edition that starts at \$25,000. See the [sample reports](#). There is also an [online demo](#).
- [ClickTracks](#) - \$495 for a standalone server and runs through a log file import. See this Builder.com article, [Get a grip on your site traffic with ClickTracks](#). A [Review](#) by Amy Hooker on Sitepoint.com. Clickz.com gave Clicktracks the [Best Web Site Analytics Tool](#) award for 2003.

Outsourced

These programs reside on your Windows desktop computer and read the log files that you download from your web server. Some of the application makers also have outsourced versions so that the user does not have to worry about installation or maintenance.

- NetIQ's [WebTrends](#). See this eWeek [review of WebTrends Reporting Center](#), the high-end product. A more conventional [Log Analyzer](#) requires the insertion of a JavaScript code snippet in each web page.
- MediaHouse [DeepMatrix LiveStats 6.2 Corporate Version](#) \$695

Maintained by [John Pike](#)

Last Modified: March 26, 2004 - 20:45

[Copyright](#) © 2000-2005 GlobalSecurity.org All Rights Reserved