Surfstats

Website Traffic Analyzer

User Manual

Table of contents

Surfstats Website Traffic Analyzer

SurfStats Website Traffic Analyzer in general allows you to analyze:

- The origin of visitors, be it from search engines, other sites or directly typed URLs.
- Classification of visitors as humans or spiders.
- The behavior of visitors and their interaction with your site content.
- Site Content Effectiveness.
- Traffic and Volume Reports

As source of information SurfStats processes the log files generated by the website and generates reports based on the information extracted from the log files.

Origin of Visitors
It is useful to know where visitors come from as you can target your marketing activities based on this information. The Search Engine Results reports will tell you a lot about the origin of your site visitors. These reports base their results on referrer information present in the web server log files.

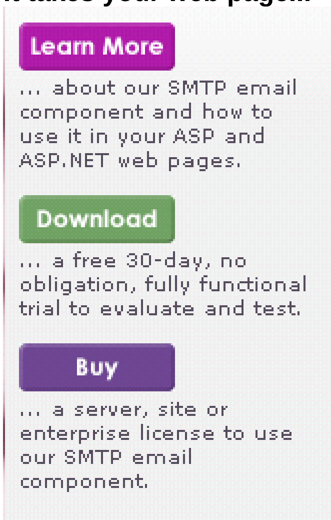
Classification of Visitors as humans or spiders
The Visitors and Spiders report categories reports separately on Visitors and Spiders. Spiders (or robots) are sent from search engines to inspect your site content. The Visitor and Spiders reports help in establishing basic visitor information, for instance the countries they originate from, how many of them return, how long they stayed etc. Spiders in general follow a fixed path through a site based on the order of the hyperlinks while humans in general follow a random path driven by content of interest.

Visitor Behavior and Interaction with Site Content
Knowing how visitors behave on your web site, which pages they look at, what files they download, how long they stay on specific pages, what keywords brought them to your site, allows you to optimize your site to maximize the benefits available to you.

People visiting your site all take different paths through your site. Surfstats Website Traffic Analyzer allows you to navigate through your site while seeing which links your visitors clicked on to move through your site.

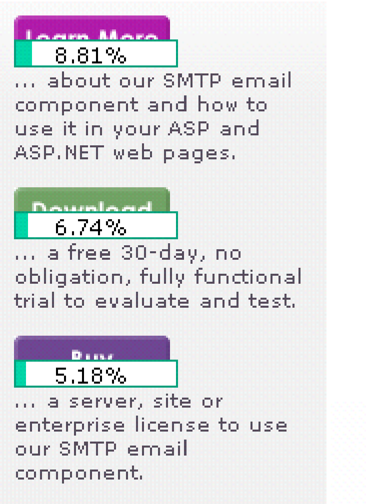
SurfStats overlays details of visitor behavior directly onto your web pages in an interactive way in the Visitor Navigation report available in the Data Mining reporting section after report generation..

It takes your web page...

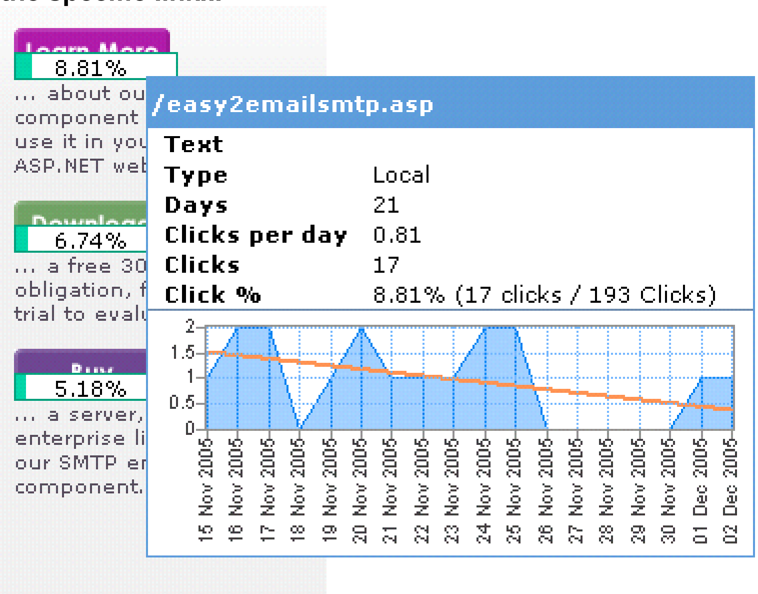


inspects the contents of the web page for links....

...and overlays it with the navigation details...



SurfStats even takes it a step further.. By hovering the mouse over a link you can see **the trend for the specific link...**



Every link on your page can be followed. SurfStats can even track links going to external sites by placing javascript into your web page, which first record the hit of any external link after which the browser is redirected. By following the links yourself you can see how your visitors navigated through your site and how long they stayed on specific pages.

To find out more about the behavior of visitors to your site please refer to the reports in the Data Mining, Visitors, Conversion Tracking and Scenario Analysis section.

Site Content Effectiveness
Site Content is the material on your site that attracts visitors, which include the pages, files, audio files, video files etc. The reports in the Page Reports category will tell you which pages are the most popular, which files are the most downloaded etc.

Traffic and Volume Reports
In general you need to know how busy your site is. The Traffic Reports will provide you with the information per day, month and year on bandwidth usage etc.

Product Name Change

SurfStats Website Traffic Analyzer was previously known as **SurfStats Log Analyzer**. As of version 8.1 the product name has been changed to **SurfStats Website Traffic Analyzer**.

Product Editions

There are currently three editions of SurfStats Website Traffic Analyzer, namely:

Standard Edition - You can analyze the log files of up to 10 sites at a time per single user license.

Professional Edition - The Professional Edition has all the features of the Standard Edition plus…

- It is licensed to analyze the log files of any number of sites per single user license on one computer for the company that registered the professional edition.
- IIS Import Wizard.
- The ability to backup site profiles, schedules and global options to a single XML file.

Enterprise Edition - The Enterprise Edition has all the features of the Professional Edition plus…

- The analysis and reporting of log files generated by proxy servers, firewalls and streaming media servers. This includes a lot of new reports for analysis of what is happening on your Intranet, firewall and media server(s).
- E-Commerce reporting

System Requirements

To properly analyze web traffic to your web site with **SurfStats Website Traffic Analyzer**, the following is required:

- A Windows XP/7/8 operating system or Windows 2000/2003/2008/2012 Server. It is not a requirement for SurfStats Log Analyzer to be installed on the web server as SurfStats can retrieve log files, analyze log files and upload report files to a web server (if required). You need an operating system that supports Windows services to use the SurfStats Scheduler.

- Administrative rights on the system if using Windows 2000/2003/2008/2012 Server.

- Microsoft Internet Explorer (IE) 7 or later. Reports are generated in HTML format and can be displayed and printed by most browsers. It is recommended to always use the latest version of Internet Explorer.

- Microsoft Word/Office if you need to create reports in Word format.

- An Internet connection.

- Pentium-class PC (300 MHz or better recommended).

- Super VGA monitor with a screen resolution set to at least 800 x 600 (higher recommended) and 16bit or 32 bit color quality graphics card.

- A mouse.

- 256 megabyte (MB) of RAM) or higher as required by the Operating System.

- 30 MB disk space to run, but about 70 MB is required to complete the installation. Additional space would be required for your log files and the SurfStats databases.

- Access to a printer with printer driver installed if printing of reports is required.

- Access to the web server log files if the log files reside on another system.

- Web server log files with referrer information.

Purchasing Information

SurfStats Software uses payment contractors for ordering and payment collection. You can order secure online, by postal mail, by e-mail, by fax or by phone.

After purchase you can obtain an activation code by clicking
Help – Register SurfStats – Request Activation Code from SurfStats

After you have supplied SurfStats with your machine identification code, SurfStats will email you a license file specific for your version and the unique identification number of the computer the product is installed on. The machine identification number is generated by **SurfStats Website Traffic Analyzer**.

See the online order page for the latest purchasing information at:
http://www.surfstats.com/buy.asp

Support

Contact information

Please use the SurfStats Helpdesk at http://www.surfstats.com/helpdesk to contact us if you need any pre-sales or technical support.

Answers to Frequently Asked Questions can be found in the FAQ section of the Helpdesk.

We only provide email support and preferably through the Helpdesk. The main reason for this is that most questions cannot be answered immediately and in most cases further investigation is required.

Unfortunately there can be a numerous potential causes for the product not to work as expected. It is difficult to give specific answers to questions without sufficient details about your circumstances, hardware and software setup, operating system and log file format. Please provide us with the relevant details and include a log file if applicable so we can reproduce the error or problem.

Program Manual

The manual and help screen are available from the main menu or the toolbar or if you press F1.

The manual is also available in Microsoft Word and Adobe Acrobat format on the download page at http://www.surfstats.com/download.asp

Program Version

You can find out the **SurfStats Website Traffic Analyzer** Version and Build number from the main menu by selecting Help - About. The displayed panel will also display your machine id.

Disclaimer

Disclaimer of warranties

SurfStats Software makes no express warranties and correctness regarding the software program and the manuals or user guides (collectively referred to as the software). The software is provided "AS IS" without warranty of any kind. SurfStats Software disclaims all warranties with regard to the software; express or implied; including; without limitation; any implied warranties of fitness for a particular purpose; merchantability; merchantable quality or non-infringement of third party rights. SurfStats Software's continuous product improvement makes specifications, features and documentation subject to change without notice.

Limit of liability

In no event will SurfStats Software be liable for any loss of use; interruption of business; or any direct; indirect; special; incidental; or consequential damages of any kind (including lost profits) regardless of the form of action whether in contract; tort (including negligence); strict product liability or otherwise; even if SurfStats Software has been advised of the possibility of such damages.

Trademarks

All trademarks and registered trademarks used in this manual are either trademarks or registered trademarks of the relevant trademark owners.

License Agreement

Please refer to the **License Agreement** program item in the **SurfStats Version....** program group (Start - Programs...).

Installation

We suggest you always run a new version on a test system until you are satisfied that it works for you. Also make a backup of your system before installing the application. The new version has been thoroughly checked but it is impossible for us to check all possible system configurations.
The same setup application is used for the SurfStats Free, SurfStats Standard, SurfStats Professional and SurfStats Enterprise Edition. The Edition can be switched from within the SurfStats application (trial version only).

New Installation

Close all other applications (recommended) before installing **SurfStats Website Traffic Analyzer**. Make sure that you install it while logged in as Administrator or a member of the Administrators group if you are installing it on Windows Server.

If the application is in zipped format, use an unzip application like Winzip to unzip it first to a temporary folder. Then run the Setup.exe file to install SurfStats Log Analyzer. If the application is in the format of a self-extracting program, run the program to install SurfStats Log Analyzer.

Removing the Scheduler

If you need to remove Remove the SurfStats Scheduler Service do it as follows:

- **Stop and Uninstall** the SurfStats Scheduler Service if it has been installed (remove from **Service** menu item in **Schedule Tasks**)
- Exit the Scheduler.

The SurfStats Scheduler Service can also be stopped from Control Panel – Administrative Tools – Services. Please note that from V7.0 the SurfStats scheduling is done with three components:

1. The SurfStats Scheduler Service which controls the Scheduler Runtime application.

2. The SurfStats Scheduler (within the SurfStats Website Traffic Analyzer Application) which is used to set up scheduled tasks.

3. The Scheduler Runtime application which executes the schedules. Please refer to the **Scheduling tasks** section for more detail.

Installing over a previous version

If you are running an older version of SurfStats Log Analyzer or Website Traffic Analyzer and you want to install the latest SurfStats Website Traffic Analyzer you have to uninstall the older version AND delete the \Program Files\SurfStats directory as the database structure between the two versions would be different.

You can save and restore your profiles and schedules (version dependant) **if** you are running the Professional or Enterprise Edition. The profiles and schedules can be restored, but

Close all running applications and make sure that you are logged in as Administrator or a member of the Administrators group if you are installing it on a Windows Server.

If you are running SurfStats Log Analyzer V6.6 or later, but pre-V7.0 click on **File – Save Profiles** to export your site profiles to a file with a name of your choice to use later.
If you are running an older version of SurfStats Log Analyzer (V6.5 beta or earlier, but V6.0.1.1 or later) make a copy of the surfsite file in the \Windows or Winnt directory.
If you are running a version prior to V6.0.1.1 you would have to recreate your site profiles after installing V7.0 or later.

If you are running V7.0 or later you can perform a backup of your current **profiles, schedules and global options** with **File - Backup**. Please refer to the section **Saving and Restoring Data** for more information on this. This option is not available in the Standard Edition.

Remove the SurfStats Scheduler Service (important) as described above in the section **Removing the Scheduler** if it has not already been done in the section **Uninstalling**.

Use the **Add/Remove Programs** from the Windows Control Panel to remove the currently installed version of SurfStats Log Analyzer from your hard disk if it has not already been done in the Uninstalling section.

Delete the old SurfStats installation folder (normally c:\Program Files\SurfStats or c:\Program Files\SurfStats<Version>) with Windows Explorer.

If the program is in zipped format, use an unzip application like Winzip to unzip it first to a temporary folder. Then run the Setup.exe file to install SurfStats Website Traffic Analyzer. If the program is a single self-extracting program, run the program to install SurfStats Website Traffic Analyzer. If you are prompted to overwrite the existing **SurfStats Log Analyzer.msi** file select Yes and follow the instructions to complete the installation.

If you were previously running a version prior to V7.0 restore the site profiles from **File - Restore – Pre Version 7.0** and select the surfsite file (or the file that you backed up your profiles to), then click the Open button. Rebuild the profile database (by clicking the **Rebuild Database** menu item on the left and selecting Rebuild All).

If you were previously running V7.0 or later and you have a backup in XML format created by SurfStats Log Analyzer V7.0 or later, you can restore the backup by selecting **File – Restore – Version 7.0 onwards**. Please refer to the section **Saving and Restoring Data** for more information.

Unfortunately we cannot guarantee that old profile file formats will be properly restored due to the changes in the file format. Please note that when you install SurfStats some sample site profiles are provided by default. If you restore your site profiles you will not be able to access these samples. If you want to have a look at these samples it is recommended that you do a **File - Backup** before restoring your profiles, this will store a copy of the profiles that comes with the SurfStats installation, for you to restore later.

Uninstalling

If you have an older version of SurfStats Log Analyzer installed or you have problems installing SurfStats, remove the program completely before installing the new version. Perform the following steps to do this:

1. Make a backup of your site profiles by selecting **File – Save Profiles** (Versions prior to V7.0) or **File – Backup** (V7.0 and later).

2. Remove the Scheduler Service (if it is running) as described under the section **Removing the Scheduler**.

3. Use **Add/Remove Programs** from the Windows **Control Panel** to remove the currently installed version of SurfStats Log Analyzer or SurfStats Website Traffic Analyzer from your hard disk.

4. Run the Registry Editor (Start - Run - regedit<Enter>) and delete the key HKEY_LOCAL_MACHINE\SOFTWARE\SurfStats.

5. Delete the \Program Files\SurfStats and \Program Files\Common Files\Borland Shared directories (**only** versions of SurfStats Log Analyzer **before** V7.0). Older versions of the BDE stored the BDE files in a different location. If the directory \Program Files\Common Files\Borland Shared does not exist find the file idapi32.dfg and delete the directory in which this file is located. V7.0 and later of SurfStats Log Analyzer and SurfStats Website Traffic Analyzer does not make use of the BDE anymore.

6. Delete the SurfStats installation folder with Windows Explorer (normally c:\Program Files\SurfStats

or c:\Program Files\SurfStats<version> [for instance c:\Program Files\SurfStats7100] if V7.1 or later).

Tip: We recommend that you use the excellent Revo Uninstaller from http://www.revouninstaller.com/ to uninstall the program. It will delete the program, directories and registry entries quickly and easily.

Switching Editions

After you have installed SurfStats Website Traffic Analyzer, run the application and click on the **Select SurfStats Edition** menu item on the left hand menu. You can then select to run the application either as:

- SurfStats Standard
- SurfStats Professional
- SurfStats Enterprise Edition

Note that if you purchase SurfStats, the Activation file will only work on the Edition of SurfStats that you have purchased. After activation the **Select SurfStats Edition** menu item is removed from the list of menu items on the left.

Visitor Origin Database

The **Visitor Origin Database** is an optional extra add-on database that contains pre-resolved IP addresses, based on Whois queries.

A separate setup application needs to be run to install the **Visitor Origin Database.** After the installation of SurfStats Website Traffic Analyzer itself, shut down SurfStats Website Traffic Analyzer (if it is running) and stop the **SurfStats Scheduler Service** (from **Control Panel – Administrative Tools - Services**) if the service is running.

Install the **Visitor Origin Database** to the **same folder** as where SurfStats Website Traffic Analyzer is installed - normally it is
C:\Program Files\SurfStats.

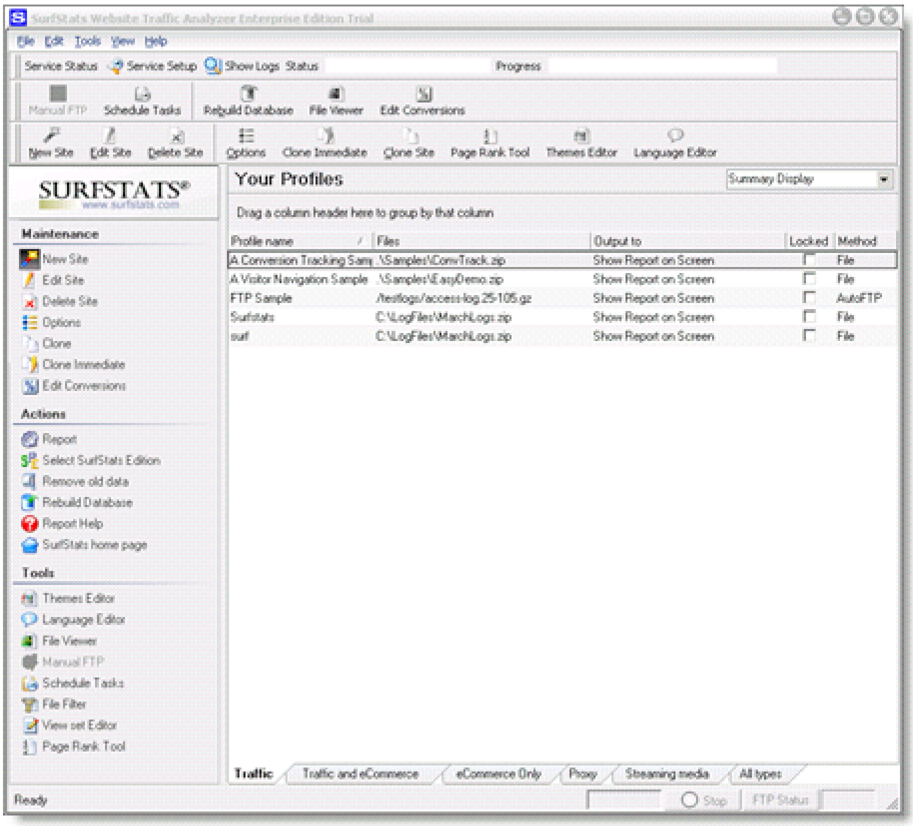
The Visitor Origin Database is registered (separately from the SurfStats Website Traffic Analyzer product). First you need to obtain an activation code by clicking **Help – Register Visitor Database – Request Visitor Database Activation Code from SurfStats**. After you have received an activation code by email from SurfStats use **Help – Register Visitor Database** to activate the database. It is important to register the database before generating reports for a site profile.

Please visit http://www.surfstats.com/sla_visitororigin.asp for more information on the **Visitor Origin Database**.

User Interface and Tools

Site Profile Mode

When you start up SurfStats Website Traffic Analyzer a screen similar to the following will display:

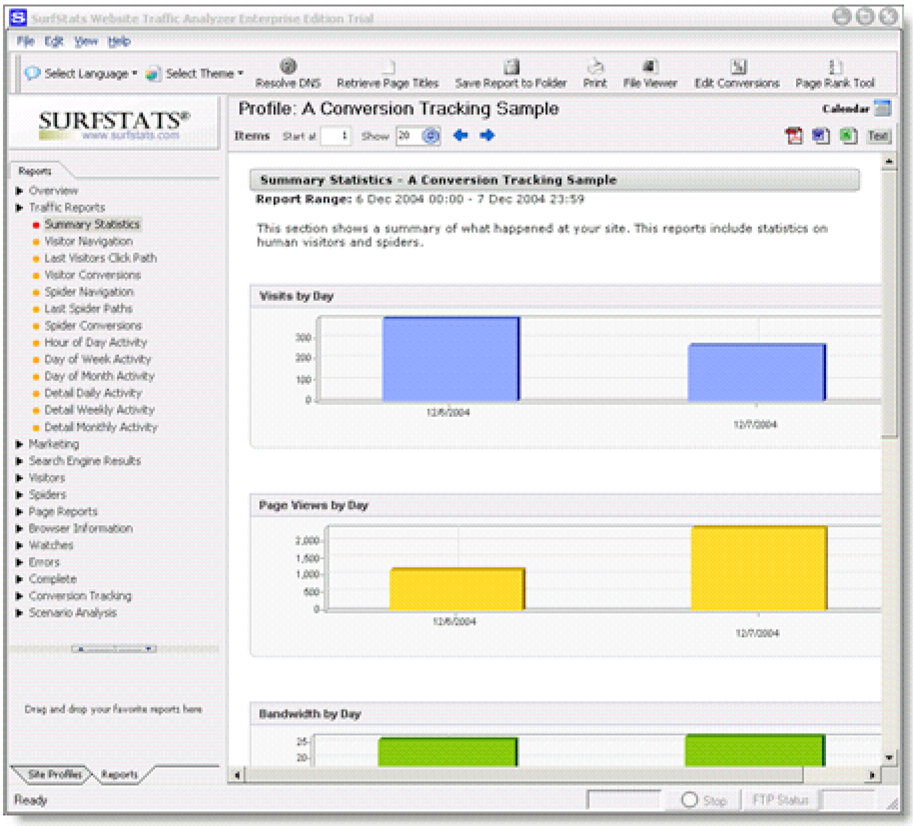


After startup SurfStats Website Traffic Analyzer is in Site Profile Mode, with only the Site Profiles tab displaying at the bottom left. The screen contains the following sub-items or elements:

- Main Menu (File, Edit, Tools, View and Help) at the top.
- Site Profile Panel on the left containing the Maintenance, Actions and Tools sections.
- Toolbar (New, edit, delete, clone profile and Log File Viewer buttons).
- Site profile list on the right-hand side.
- Status Bar at the bottom.
- Extra tabs are displayed at the bottom right when SurfStats is run as the Enterprise Edition that allows switching between Web Traffic Analysis, Proxy Analysis and Streaming media Analysis.

Reports Mode

When report(s) have been generated from a site profile (by double-clicking on the profile), the program changes from the Site Profile mode to Reporting mode. The Site Profile panel gets overlaid with the Report Panel. The items on the main menu and the toolbar also change. In Reporting mode a screen similar to the following is displayed:



In Reporting Mode the screen contains the following sub-items or elements:

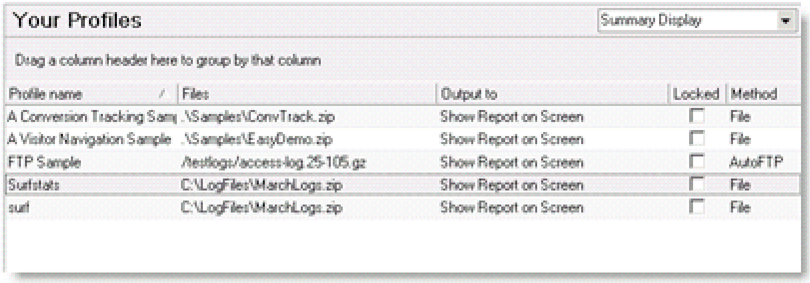
- Main Menu (File, Edit, View and Help)
- Report Options Toolbar at the top (Select Language, Select Theme etc.). The Report Options toolbar is dock-able and can be moved around.
- Schedules Toolbar at the top (Service Status, Service Setup etc). The Report Options toolbar is dock-able and can be moved around.
- Reports menu items (Overview, Traffic, Visitors, Pages etc.) on the left. The selected report will display if the report menu item is clicked on the selected report category tab.
- Favorites report group at the bottom left where you can store a list of your favorite reports, similar to favorites in a browser.
- Report Viewer on the right-hand side where the reports are displayed.
- Start at and Show selectors above the Report Viewer allowing the selection of the start report item, the number of items to display, next and previous buttons. These selectors are used to navigate through report content.
- Site Profiles and Reports panel selector for switching between Site Profile and Reporting mode (Profiles and Reports tabs at bottom left).
- Status Bar at the bottom.
- The Reporting mode panel is closed if you switch back to the Site Profile mode (by clicking on the **Site Profiles** tab).

Site Profile panel

The Site Profile panel lists all the site profiles (in Site Profile mode) and is displayed on the right hand side of the screen.

You can rearrange the columns (of the profile list on the right) by dragging and dropping the column header to a new location.

You can sort the site profiles by column header by clicking on the column header. You can change the sorting order (ascending or descending) by clicking on the column header again.
You can group by multiple columns by dragging and dropping column headers (one at a time) on the box above the column headers (the text “Drag a column header here to group by that column”). The grouping order will then be displayed in a tree directory as shown below:

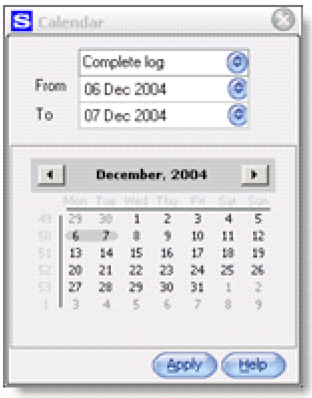


You can remove a grouping by dragging and dropping the group header from the tree back to the site profile list. You can also rearrange the grouping order by dragging and dropping it to another location.

Calendar

At the top right in **Reports Mode** the **Calendar** link menu item is displayed.

If this link is clicked the report calendar is displayed, from which you can select the date range to report on in the currently selected report.

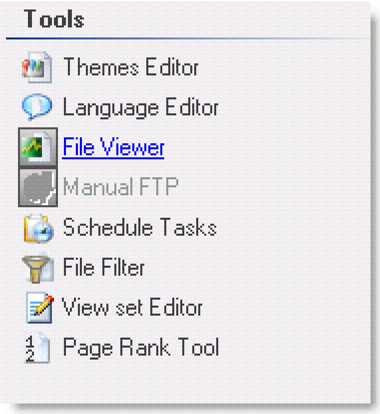


Dates can be selected on the calendar in a number of ways, as follows:

- The period selection dropdown box (displaying **Complete Log** by default) at the top of the calendar allows you to specify exact periods, for instance the complete log, today, yesterday, this week etc. If you select **Specify Period** here the From and To Date select boxes become active allowing you to specify a start and end date to report on. Click **Apply** afterwards.
- Select the first date (from the days displayed), keep the left mouse button down and move the mouse to the last date, then release the mouse button.
- Shift-click the dates.
- Clicking the left (<) or right (>) arrows next to the currently selected month name moves the period forward or backward one month relative to the current month displayed. This happens if no date on the calendar has focus. If a specific day is selected in the calendar clicking these buttons moves the selected day one month back/forward.
- Clicking the double-left (<<) or double-right (>>) arrows moves the date selection back/forward one year.

Tools

The following tools are supplied with SurfStats Website Traffic Analyzer and are available from the Tools menu at the bottom left or from the Toolbar on the left when the main window is displayed in Site Profile mode.



- **Themes Editor** - You use this tool to create or edit themes. This tool was formerly known as the SurfStats Artist. Themes are linked to a profile and determine the way reports are displayed for a specific profile (Edit profile – Appearance - Theme). Note that you do not use the Themes Editor to select a theme for a profile, but only to edit it. You select a theme by clicking on the Theme menu item, which will bring up a theme. The theme can then be edited (by selecting elements from the Item select box) and saved (by clicking Save as…).

- **Language editor** - You use the language editor to modify phrases and text on reports. You can also add new languages. Note that you can set up a report language for each site profile (Edit profile – General tab). The Language Editor can be used to create/edit Western European, Eastern European, Russian and Asian Languages. Please note that to display some languages you have to enable it from Control Panel – Regional Options. Please refer to the section Site Traffic Reports for more on the Language Editor.

- **File viewer** - You use the file viewer to inspect or view log or other files. This file viewer can also be used as an editor to correct errors in log files. The tree view on the left hand side is used to browse to the drive and folder of the file to view or edit. The top panel on the right hand side is used to select the file for display. If it is a compressed file, the contents of the compressed file will be displayed in the middle panel. The bottom panel will display the selected file. If the displayed file is a text file, you can edit it, find selected text or highlight phrases in it.

- **Manual FTP** - The Manual FTP client is used to manually transfer files between FTP servers on the Internet/Intranet and your computer. You would normally use Manual FTP to download log files from a host FTP server. Note that alternately you can set up automatic FTP for every site profile. This would initiate automatic downloading of files from a FTP server. The Manual FTP menu item only becomes enabled when a profile of one of the FTP types is selected.

- **Schedule Tasks** - You use this tool to set up tasks to be executed at pre-determined intervals, e.g. to retrieve log files every morning, followed by analyses and reporting.

- **File Filter** – A tool to remove unwanted log entries from log files before processing them with SurfStats Website Traffic Analyzer.

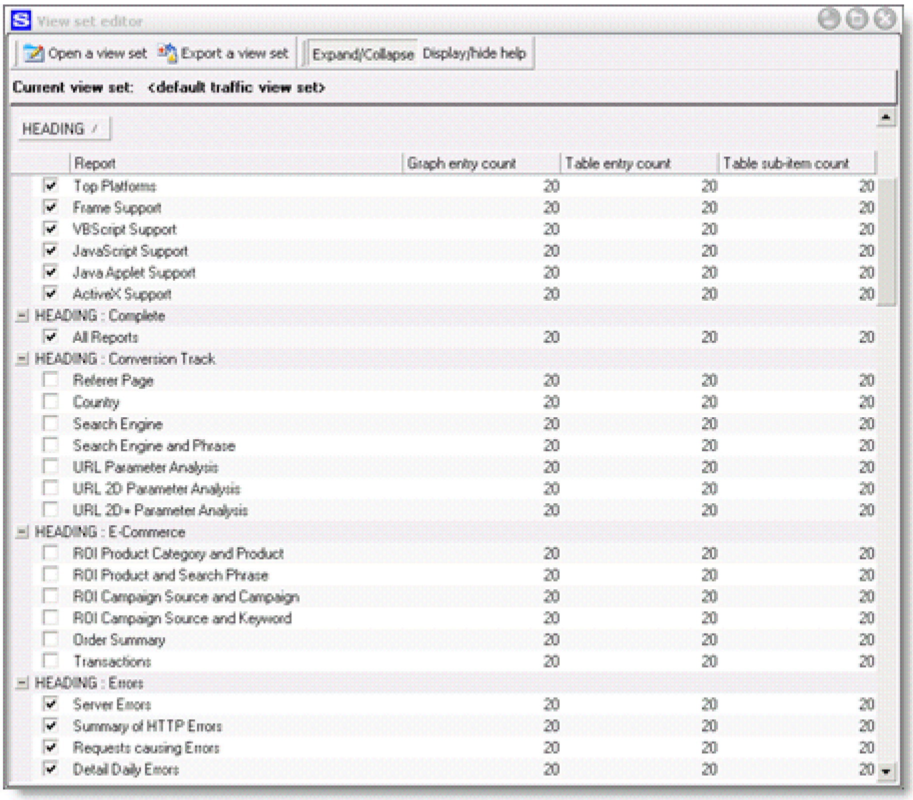
- **View set Editor** – Use this tool (available in Professional and Enterprise editions) to change the views information for multiple profiles. More information is available in the View set Editor section.

- **Page Rank Tool -** Use this tool to determine the page rank for your site at search engines for specific keywords.

View set Editor

The **View Set Editor** (available in Professional and Enterprise editions) allows a specific view set to be edited and simultaneously applied to one or more site profiles. A view set is a template indicating for each report how many graph, table item and table sub-items to display. Please note that the settings here **does not affect reports generated to screen**, only to folder, FTP and email. An alternative method for editing view information is to click **Views** on the left when editing a site profile, but this allows only editing the view information for the profile being edited.

To access the **View set Editor** click the **View set Editor** menu item on the left in the **Tools** section. The following screen is displayed:



Click the **Open a view set** button to display the currently available view sets. There is a view set for each profile defined as well as the three types of profiles: traffic, proxy and streaming media. The proxy and streaming media view sets is only applicable to these types of profiles in SurfStats Website Traffic Analyzer Enterprise Edition.

After a view set has been selected the view set can be edited by selecting/unselecting a report or specifying the number of graph and table entries. The changes are stored automatically to the currently selected view set. To apply the view set click the **Export a view set** button. A dialog displaying a list of profiles is displayed. Select the profiles to apply the view set to and click the **Export** button. The view information as well as the view sets for the selected profiles will be updated.

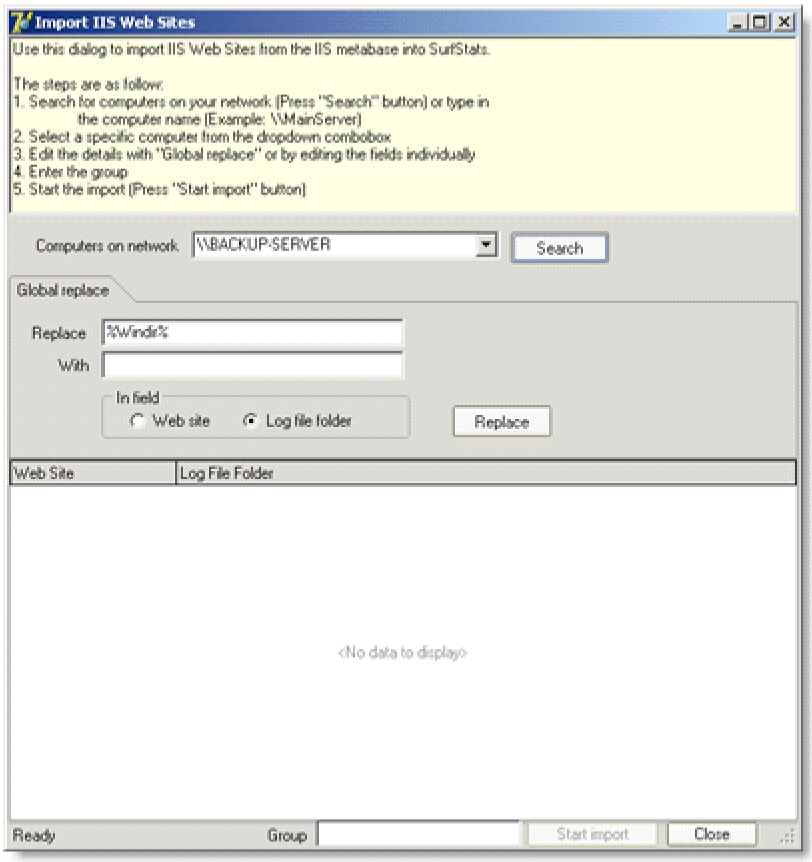
Note that if a new site profile is created the view information from the selected profile type will be used. For instance if a new profile of type traffic is created the view set information from <default traffic view set> will be used.

To change the view set information for a specific profile open the view set for the profile with the **Open a view set** button then change the settings as required. The changes will automatically be applied. The **Expand/Collapse** button expands/collapses the list of reports for the types of reports. The **Display/hide help** button display/hide the help on how to use the **View set Editor** at the top of the screen.

Importing IIS Web Sites

This feature is only available in the Professional and Enterprise editions of SurfStats Website Traffic. Click on File – Import IIS Wizard to import IIS Web Sites from the Windows IIS Metabase:

The following screen should appear:



Follow the steps 1 to 5 to import the web sites.

Please note: A remote server can only be accessed if IIS is installed on the remote server. You also need the Microsoft ADSI client installed on the computer running SurfStats Website Traffic Analyzer, which is installed by default on Windows 2000 and later Operating Systems. Please visit www.microsoft.com/adsi if you need to install the ADSI client.

Changing number to report on

The default number of records in reports is set to 20. The number of records on graphs can be set to a maximum of 20 but the record numbers of main and sub tables can be changed. You can set the number of records to report on per site profile. There are two places where you can set the number of records to report on:

1. **Before report generation** - Select the site profile from the main screen, click on **Edit Site Profile** and then on **Views**. Set the number of Main Table Entries and click on the icon next to it. This setting is for when reports should be generated to a folder. The number can be set for individual reports.

2. **After report generation** – Generate reports and select the number of records to display in the **Show** field, top right on some of the report category tab. This setting is used when generating reports to screen. The last setting will be used when generating a new report.

Administering Site Profiles

What is a Site Profile

A site profile provides essential information about your website to SurfStats Website Traffic Analyzer. Site profiles are stored in the SurfStats main database and each profile has a separate reporting database built from the analysis of log files. Some of the information that defines a site profile is listed below:

- Site Profile Name and description
- Log file retrieval method (File, AutoFTP, Manual FTP)
- Details of the retrieval method.
- The first and last date of the log file entries (after a database population)
- The files and directory path of the log files.
- Method of database population
- Theme
- Log File and Date Format fields
- Single or Multi File Logs
- Output Destination
- Details of the Output Destination
- Logo information on reports
- Language on reports
- Image type used on reports
- Automatics Reverse DNS lookups or not
- Number of records to show on reports
- Time Zone
- Separation of Parameterized Queries
- The Default Page
- etc.

Keep a backup copy of your site profiles in case of a system crash or other form of disaster. You create a backup copy by clicking File – Backup in site profile mode. The backup option is not available in the Standard Edition.

Adding a new Site Profile

When you click on the **New Site** menu item, the New Site Profile wizard will guide you through the steps to set up a new site profile. This specifies the log files that SurfStats should use, how the files should be processed and other options for the site profile.

Click the **Forward** button and a screen will be available where you can then enter the following information:

Site Name - A physical directory with this name will be created. This is where the site's data will be stored. The physical directory will get created under the SurfStats installation directory. You can only use the standard characters that are allowed in directory names.

URL - The web page URL for the site, for instance http://www.surfstats.com

Method of log file retrieval - Select how the log files for the web server will be accessed or retrieved. Please refer to Specifying log files for more information.

Click the Forward button to progress to the wizard's next screen. Depending on the retrieval method selected the wizard will decide which screen to display. The following assumes that File was selected:

Select the log file names - You can select multiple files and directories.
Note that the first line should contain a log path and file specification. There should not be empty lines between each path and file specification.

Select the Profile Type. Only Traffic is available in the Standard and Professional Edition. Traffic, eCommerce, Proxy and Media are available in the Enterprise Edition. Different types of reports get generated depending on the profile type selected. If you select Traffic the new profile will be created on the list on the Web Traffic tab. If you select Proxy the new profile will be created on the list on the Proxy tab. If you select Media the new profile will be created on the list on the Streaming media tab. This can be changed later by editing the site profile. In the Enterprise Edition if you created the profile from the Web Traffic Analysis tab, Traffic will be selected by default here. If you created the profile from the Proxy tab, Proxy will be selected by default, the same with Media.

Please use different site profile names if you want to analyze a site in more than one way, for instance if you want to analyse the same site as a traffic site and a proxy site use different site profile names.

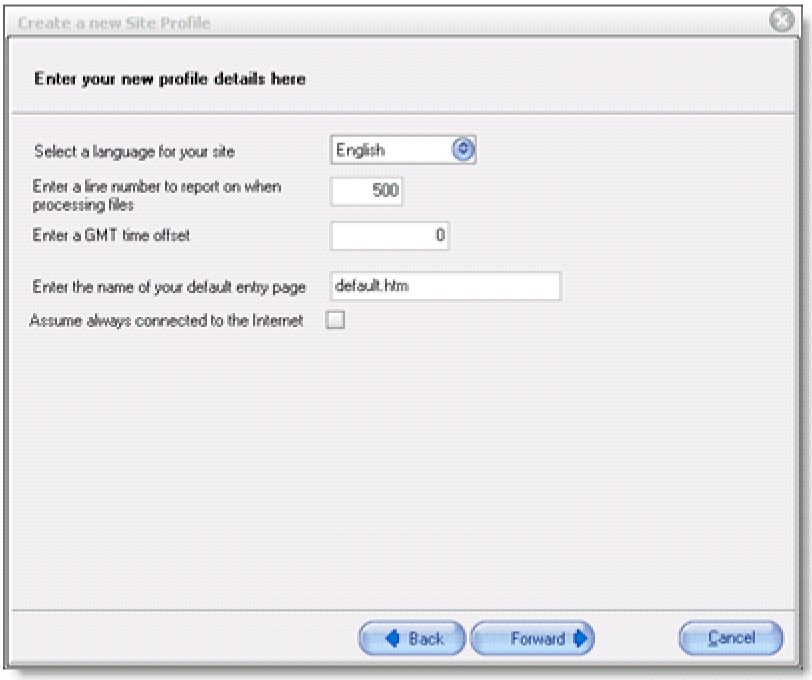
Select the log file format - Use Auto-detect whenever possible.

Select the format of dates within the log files - Use Auto-detect whenever possible. If you have problems analyzing log files from Apache web servers inspect one of the log files with a text editor and specify the format explicitly, for instance MDY.

Click the Forward button to progress to the next screen.

Select the report destination, that is the destination where you want reports to be generated to. Depending on the selection different fields will become available.

Click the Forward button to progress to the next screen.



Select the language for your site, the number of lines to read before processing, the GMT offset, the name of your web site's default entry page (as set up on the web server) and if the system has a permanent connection to the Internet. Use the defaults when in doubt.

Click on Forward and click Finish to complete the wizard steps.

To start report generation for a profile double click on the newly created site profile or select it and click on Generate report to retrieve the log files and generate reports for the specific site profile.

For more specific information on profile fields please refer to the Editing a Site Profile section.

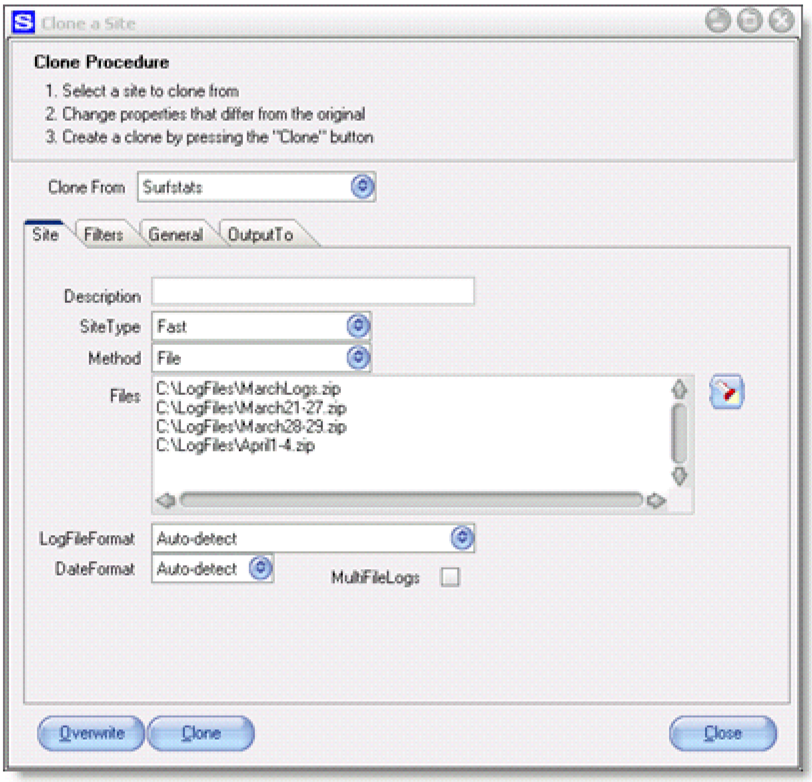
Deleting a site profile

To delete a site profile, select the site and click on the **Delete Site** menu item.

Site Profile Cloning

This option in version 7 and later is available in the Standard, Professional and Enterprise Editions of SurfStats.

The Clone feature is used to duplicate existing Site Profiles quickly without the need to re- enter fields that should contain the same values as the profile being cloned. To clone a site profile, select the site profile to clone and click on the Clone Site Profile menu item on the left. A screen similar to the following will appear:



Change the fields that need to be changed. If you click on the Overwrite button, the current Site Profile will be overwritten with the changed fields. If you click on the Clone button, a new Site Profile will be created and you will be prompted for a Site Profile name. Remember that the same restrictions apply for the profile name as using the Create profile wizard. You can create more site profiles from the same window by clicking on the Clone button again. Click the Close button once you are done with the cloning process.

You can also right click on a site profile and select to:

- Clone profile – This provides an option box where you can change the site profile details
- Clone immediate – This clones the site (as is) and only provides the option to change the site profile name.

Site Profile Types

Different profile types in SurfStats Website Traffic Analyzer can be selected when a site profile is created. Different reports will be generated depending on the profile type selected.

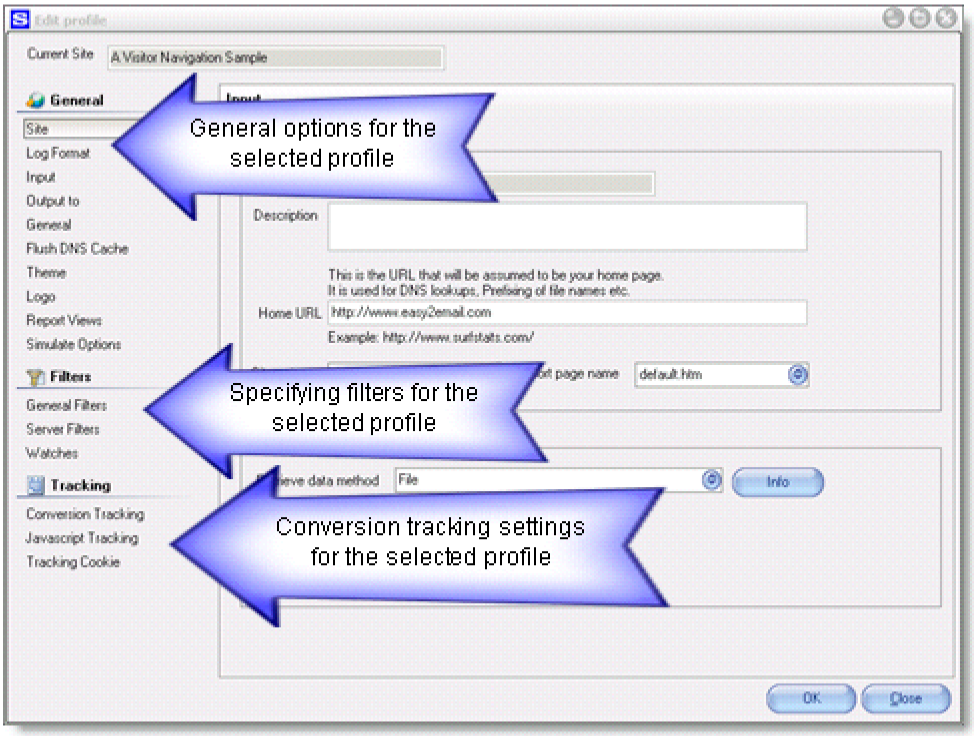
When SurfStats Website Traffic Analyzer is run as an Enterprise Edition six tabs, which can be individually selected, are visible at the bottom:

1. **All types** – Displays all the profiles of all types that were created in SurfStats Log Analyzer.
2. **Traffic** – Displays profiles of web sites associated with log files from common web servers, like IIS, Apache etc. Log analysis of these types of profiles are concerned mostly with outside visitors accessing the web site.
3. **eCommerce** – Displays eCommerce profiles. With eCommerce reports extra reports are generated for PPC campaigns, opt-ins and shopping carts.
4. **Traffic and eCommerce** – Displays traffic and eCommerce profiles.
5. **Proxy** – Displays profiles of proxy servers and firewalls and associated log files. Proxy servers and firewalls are special types of servers that are installed between an internal network and an external network (like the Internet). Log analysis of these types of profiles are concerned mostly with visitors on the inside of the network (mostly employees of an organization) accessing web sites on external networks.
6. **Streaming media** – Displays profiles of streaming media servers and associated log files. A streaming media server provides multimedia and other content to users accessing the media server with a media player (like Windows Media Player, Real media player etc.).

When a specific tab is selected and the New Site Profile wizard is run a profile of the specific type will be created. Running the New Site Profile wizard from the All types tab sets the type to Web Traffic. The type can still be changed in one of the steps in the wizard. The reports created by SurfStats differ for the different types of profiles.

Editing a Site Profile

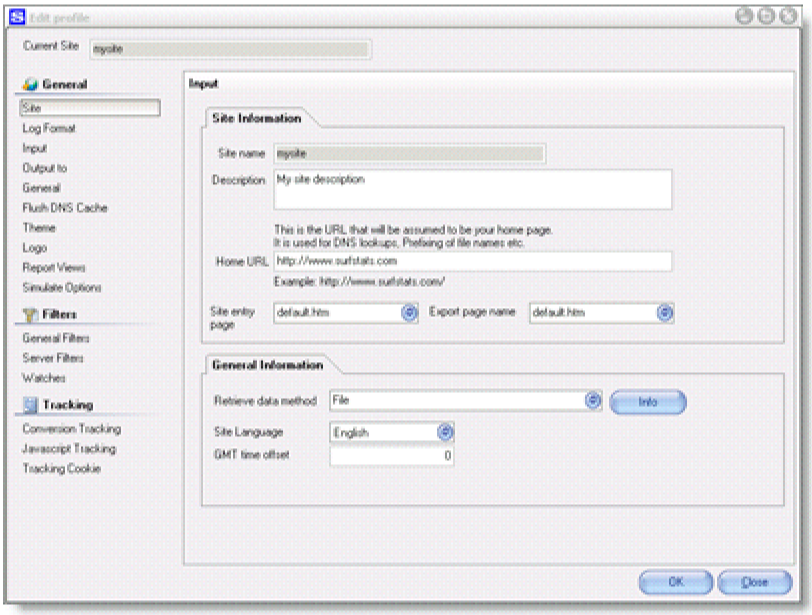
You can change the site profile settings by selecting the profile from the list of profiles and then clicking the **Edit Site** menu item on the left.



You can select the appropriate menu item on the left under the headings General, Filters and Tracking to edit the specific aspect of the selected profile.

Editing Profile Options (Site)

Select a profile and click the **Edit Site** link on the left, then click the **Site** link. Here you can change the following settings for the selected site profile:



Description - A description that you want to associate with the site profile.

Home URL - Used to specify the web site URL to add in front of the filename part for instance in the Top Pages and Most Downloaded Files reports. The web server logs only the filename part of the URL. Specifying a Home URL will ensure that links on reports can be clicked to visit the actual web page. Setting the home page is also required to enable page title lookups (as SurfStats needs to access the web site and download the web page to determine the title of the page).

Site entry page - The name of the default page on your website.

Export page name - The name of the default page in the folder where reports will get exported to.

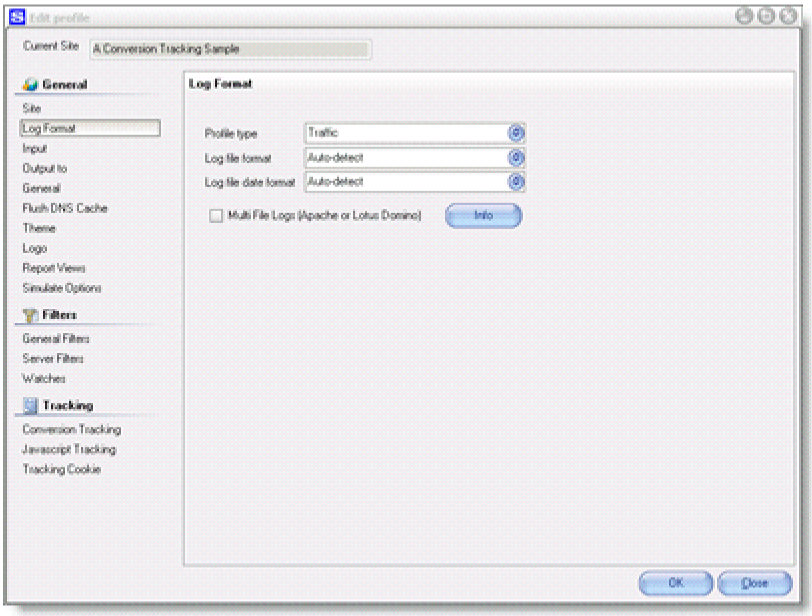
Method of Retrieval - The method that will be used to retrieve log files when reports are generated. Please refer to Specifying log files for more information.

Site language - The language that SurfStats should use when generating reports. Note that you can add or modify the report text or phrases to any Western or East Asian language with the Language Editor. Please note that if you are specifying an East Asian language that the language has to be installed on the computer system on which SurfStats is running (set up in Regional Options in Control Panel).

GMT offset - SurfStats will adjust the time zone of log entries before adding them to the specific site profile database, which will also take the server's time zone in consideration. For instance if - 5 is specified and the web server is in a time zone of GMT-05:00 for instance the following will happen for a log entry that happens at 13:37 (server time) on a Windows IIS web server: IIS will adjust the timestamp of the log entry (hit) to 18:37 (that is 13:37 – (-05:00)) and then add it to the log file. When SurfStats processes the log entry it will adjust the time of the entry added to the site profile database to 13:37 (that is 08:37 + (-5)). When creating a new profile, for log files from an IIS web server, it is suggested to use the offset as specified on the web server the log files were retrieved from, for instance if the server time is set up for GMT-05:00, then –5 will be the best choice for this field, as it will indicate the time on the server when the log entry was made. Note that the nature of your site will also influence the value you should specify here, for instance (in the sample described above) if you host your site in the US and you and most of your site visitors are from France you would specify +6 here (so 1-(-5)). On an Apache (and most Unix) web servers specify 0, as most Unix web servers do not adjust the time of the log entries.

Editing Profile Options (Log Format)

Select a profile and click the **Edit Site** link on the left, then click the **Log Format** link. Here you can change the following settings for the selected site profile:



Depending on the edition of SurfStats Website Traffic Analyzer you are running the following **profile types** will be available:

Standard Edition - Traffic
Professional Edition - Traffic and eCommerce and eCommerce Only
Enterprise Edition - Trafffic, Traffic and eCommerce, eCommerce Only, Proxy and Media

Please refer to the Site Profile Types section for more information on profile types.

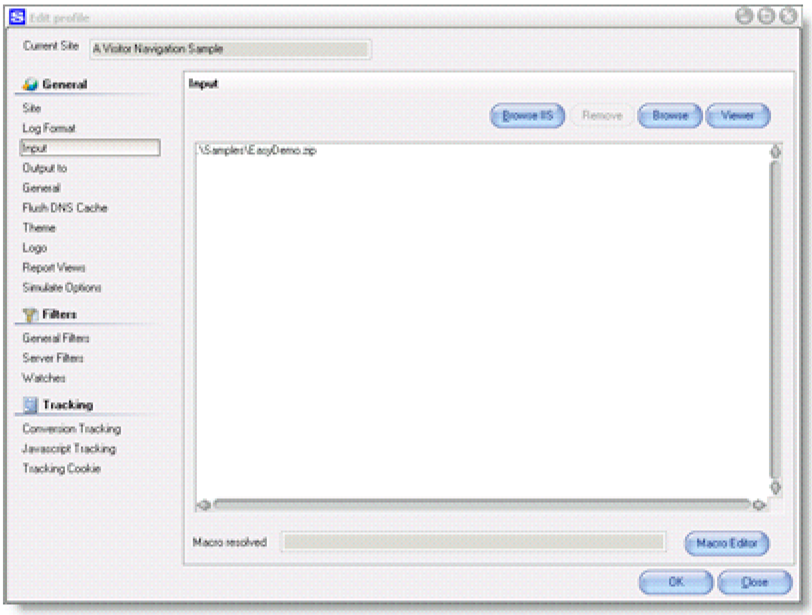
Log file format - The type of log files that this profile will analyze, for instance Apache Common, Microsoft IIS etc. For Log file format use Auto-detect as far as possible.

Log file date format - The format of the date/time field present in the log file. For Apache log files it is best to specify the log file date format explicitly.

Multi File Logs - Sometimes a web server generates multiple log files simultaneously, for instance one file contains info on files accessed, another file contains the referrer info etc. For these type of log files check the Multi File Logs... option.

Editing Profile Options (Input)

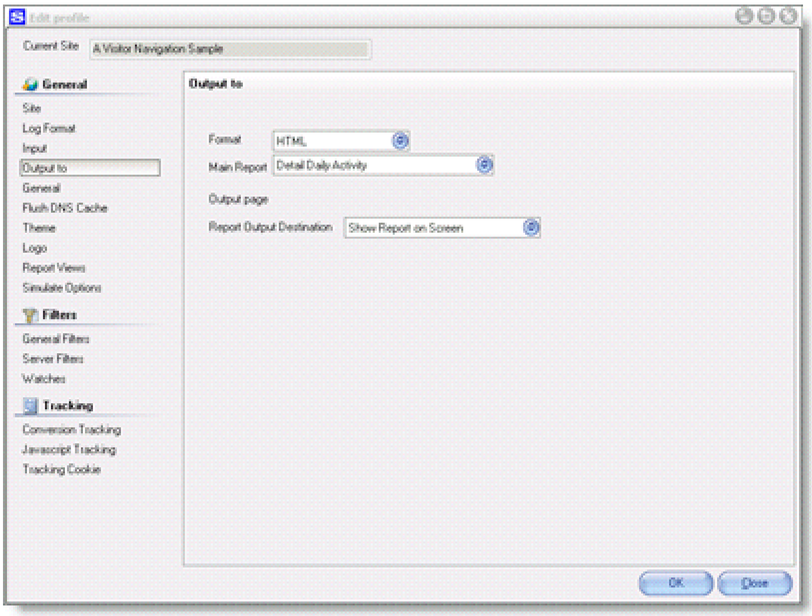
Select a profile and click the **Edit Site** link on the left, then click the **Input** link. Here you can specify the log files that should be processed to generate reports for the selected profile:



You can specify log files in uncompressed (for instance .log) and compressed format (for instance .zip files). It is best to specify the full path and filename. Please refer to the section on Working with Log Files for more information.

Editing Profile Options (Output To)

Select a profile and click the **Edit Site** link on the left, then click the **Output To** link. Here you can specify the report destination options:



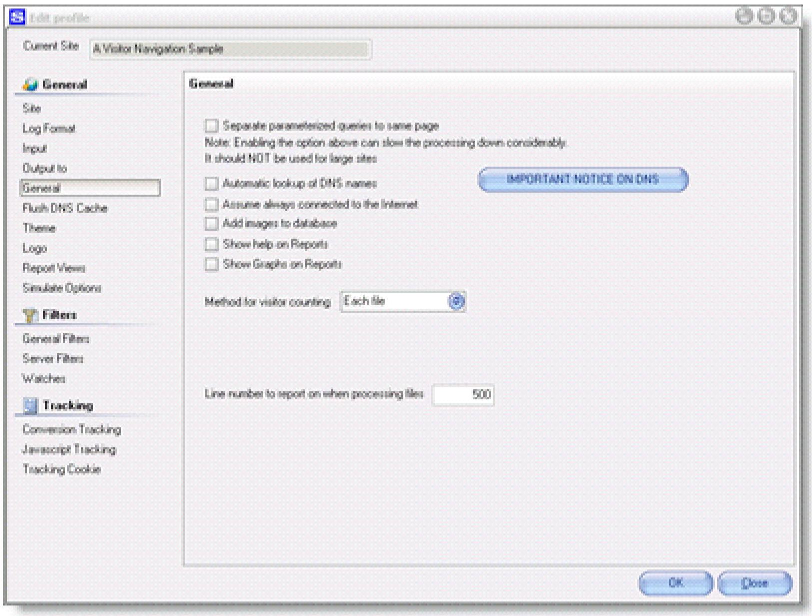
Format - You can generate reports in either HTML or MS Word format.

Main Report - Select the report that should be shown in Summary Statistics when reports for the selected site profile are generated.

Report Output Destination - The destination where reports for this profile should be generated to. Please refer to Report Destination for more information.

Editing Profile Options (General)

Select a profile and click the **Edit Site** link on the left, then click the **General** link. Here you can specify general options regarding the report generation:



Separate parameterized queries to same page - Tick this box if you want to differentiate between Active Server Pages and other pages called with different parameters, e.g. between the page "training.asp?level=1" and the page "training.asp?level=2" to show it as hits to different pages or not.

Automatic lookup of DNS names - Check this box if you want to resolve domain names (from IP addresses) automatically for the site profile after the profile's database has been updated, before report generation. Note that you have to be connected to the Internet and be able to access the DNS Server to do DNS lookups if you do not have a permanent connection to the Internet. This performs a Reverse DNS lookup of IP addresses.

Assume always connected to the Internet. Select this box if you have a permanent connection to the Internet and do not want to be prompted to be connected.

Add images to database - Tick this box if you want SurfStats to add hits to image files (like .gif, .jpg files etc.) to the site profile's database. Considerable gains can be made in report generation time if this option is not selected. Hits are classified as of type image if it is listed under SurfStats Options – File Types – ExtensionType: image.

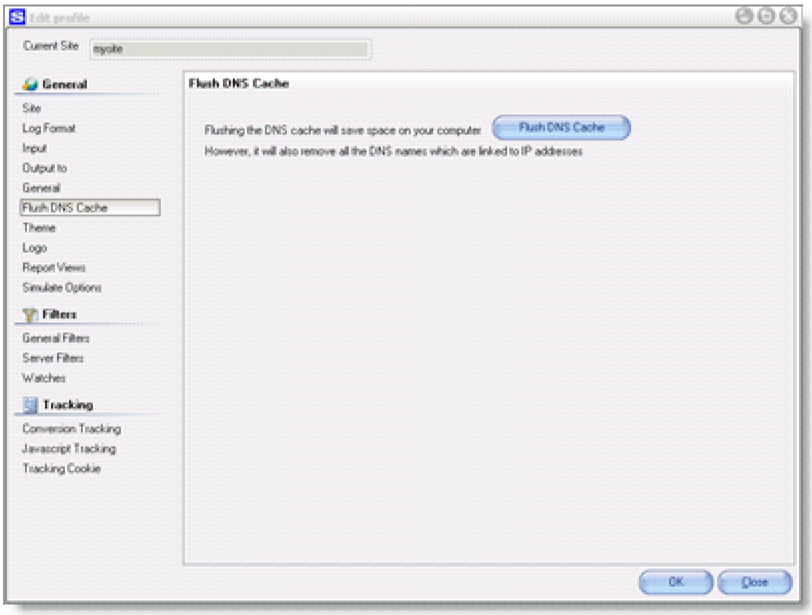
Show Help on Reports - Whether or not help information is displayed on reports.

Show Help on Reports - Whether or not graphs are displayed on reports.

Line number increment to report on when processing log files. The smaller the v
value the more frequently the progress indicator and line item counter will be updated when lines are processed by SurfStats from a log file.

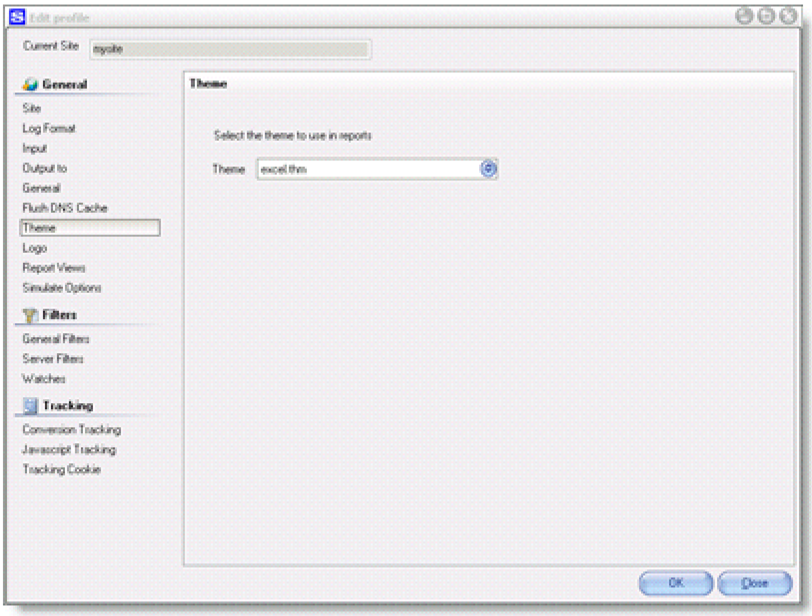
Editing Profile Options (Flush...)

Select a profile and click the **Edit Site** link on the left, then click the **Flush DNS Cache** link. Here you can clear the Resolved Reverse DNS names stored by SurfStats by clicking the Flush DNS Cache button:



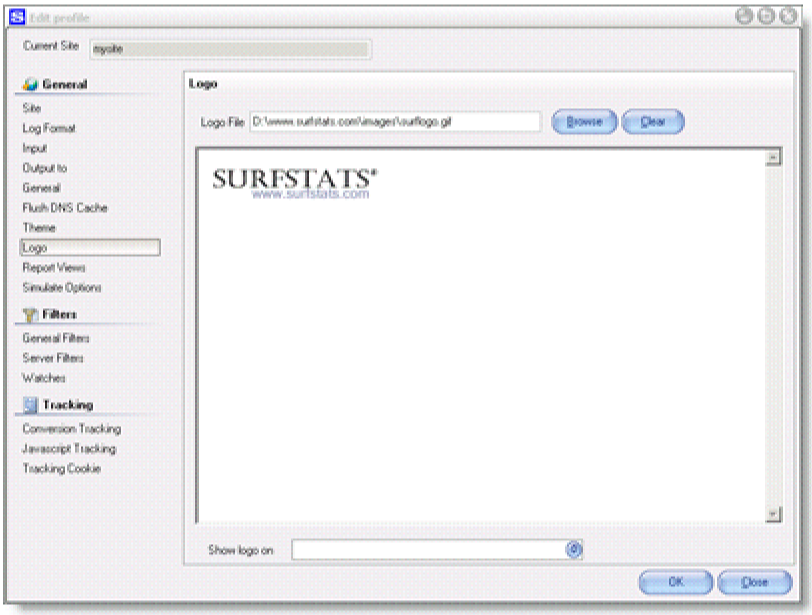
Editing Profile Options (Theme)

Select a profile and click the **Edit Site** link on the left, then click the **Theme** link. Here you can specify the theme to use on the specific profile. The theme will determine the appearance of reports.



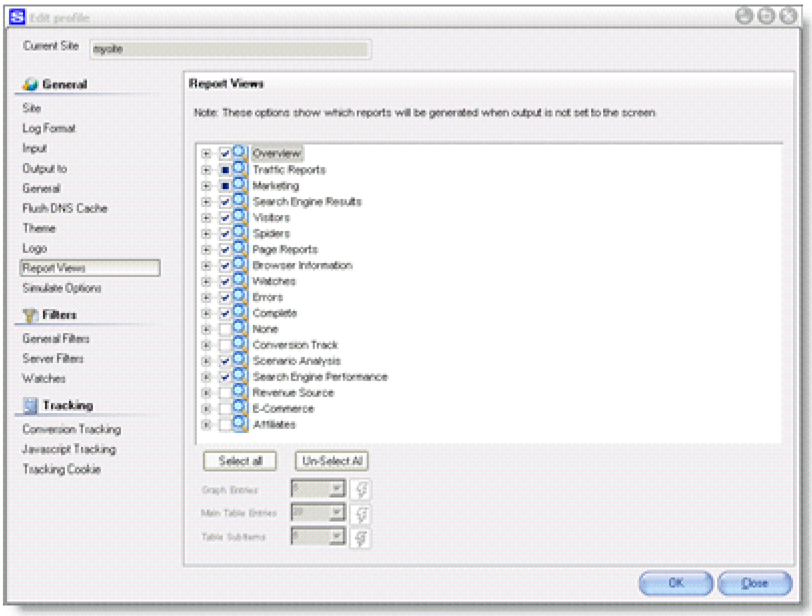
Editing Profile Options (Logo)

Select a profile and click the **Edit Site** link on the left, then click the **Logo** link. Here you can specify the logo to use on reports for the specific profile and where to use them.



Editing Profile Options (Report Views)

Select a profile and click the **Edit Site** link on the left, then click the **Report Views** link. Here you can specify which reports to generate, the number of graph entries and the number of table entries.



To change the view settings for a specific report expand the report category on the left and select the report on the right. You can:

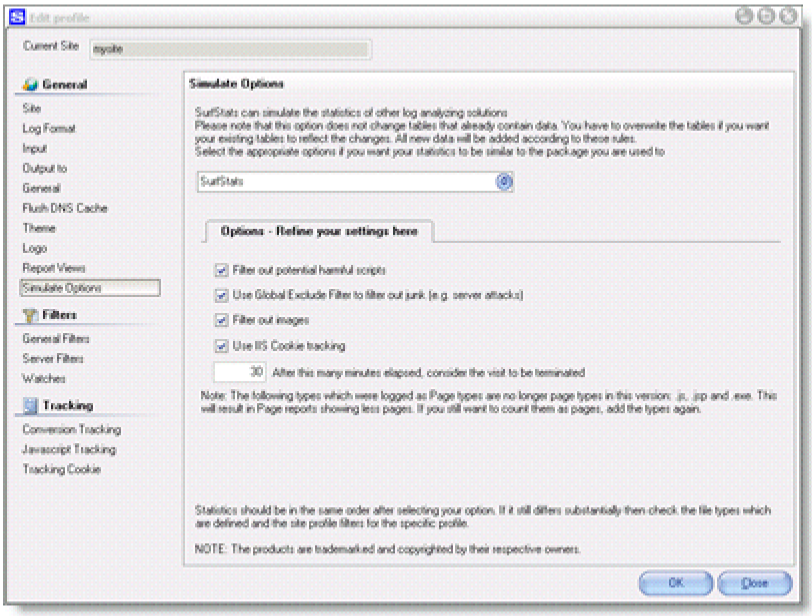
- Check/Uncheck the checkbox next to a report name to toggle report/non-report generation.
- Change the number of graph entries.
- Change the number of main table entries.
- Change the number of table subitems (if applicable)

After you have change a value click the yellow lightning button for the setting to take effect.

If you are running the Professional or Enterprise Edition of the product you can use the View Set Editor tool to change the view to apply to multiple profiles at once.

Editing Profile Options (Simulate Options)

Select a profile and click the **Edit Site** link on the left, then click the **Simulate Options** link. Here you can specify how SurfStats will generate reports, simulating a competitive product as well as log file processing options.



Package simulation - Select a competitive product from the dropdown box you want SurfStats to simulate.

Filter out potentially harmful scripts - Filters out script that can be present in log files causing harmful behaviour.

Use Global Exclude Filter.... - Check this to apply the global filter specified in Options - Global Exclude Filter. This is mainly used to filter out junk log entries caused by worms etc.

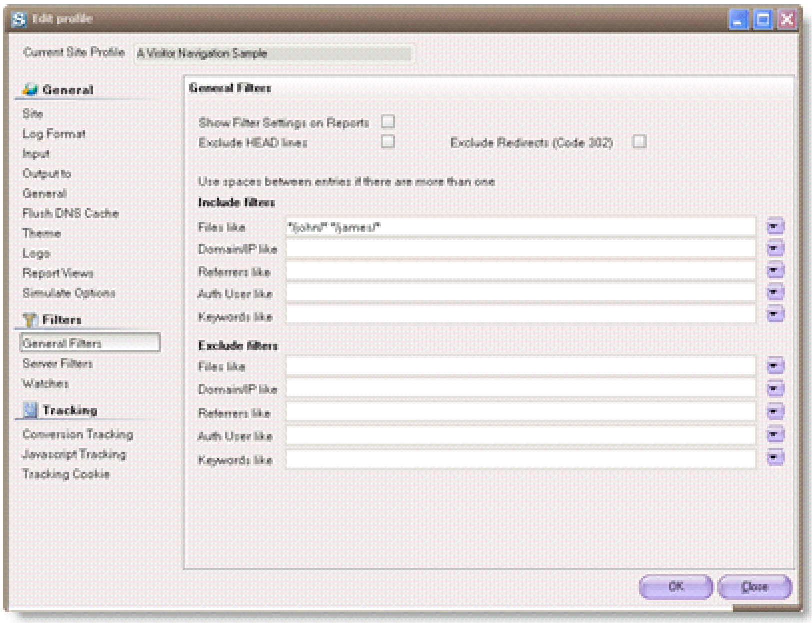
Filter out images - Omits hits for images files when processing log files.

Use IIS Cookie tracking - When processing log files, visits from different IP addresses but with the same ASP/ASP.NET session ids will be grouped together as coming from one IP address (the first found related to the cookie). This is used to provide a more accurate visitor count where an ISP use different IP addresses to service web requests from the same visitor.

After this many minutes elapsed, consider the visit to be terminated - The number of minutes between subsequent hit requests to consider hits as part of a visit or a separate visit. For instance if this value is set to 30 and hits for files from the same IP are received at 10:45AM, 10:47AM, 10:51AM, 11:30AM and 11:31AM it will be considered as 2 visits from the IP.

Editing Profile Options (Filters)

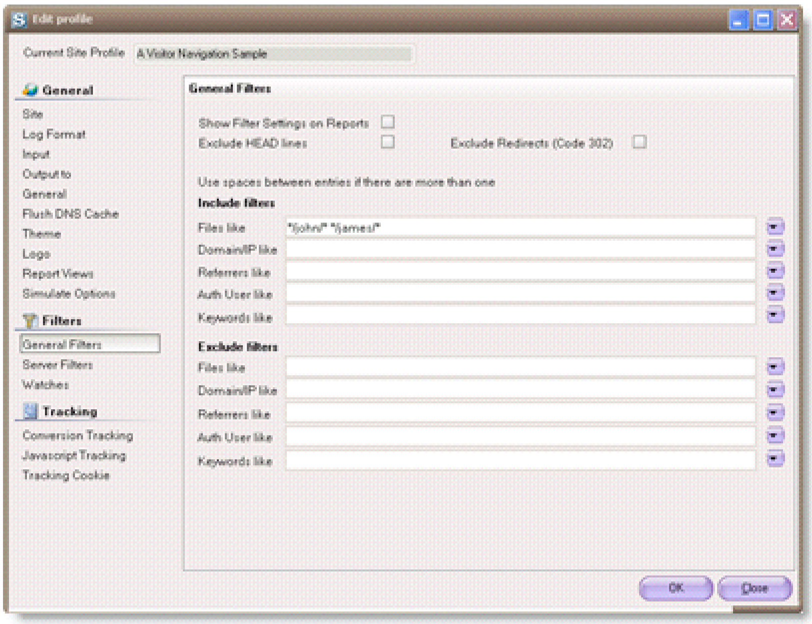
Select a profile and click the **Edit Site** link on the left, then click the **General Filters**, **Server Filters** or **Watches** link. Here you can specify filters to apply to the profile when reports are generated.



Please refer to the [Filters and Watches](#) section for more information on filters.

Editing Profile Options (Tracking)

Select a profile and click the **Edit Site** link on the left, then click the **Conversion Tracking, JavaScript Tracking** or **Tracking Cookie** link. Here you can specify conversion tracking options to apply to the profile when reports are generated.



Please refer to the [Conversion Tracking](#) section for more information on setting up conversion tracking.

Quick Start

This section will help new users to set up a basic site profile and generate reports.
To see how easy it is to generate web site traffic reports, follow the steps:

1. Set up a site profile
2. Populate the database

3. Generate reports

Setting up a site profile

When you click on the New Site Profile menu item, a wizard will lead you through the steps to create a new site profile.

We suggest you use the defaults in most cases.

On the first panel, after the wizard welcome screen, enter the site profile name, e.g. mysite, specify the URL of your website, select the File method to retrieve your log file(s) and click on the Forward button.

On the next panel, click on the Browse button, and browse to the directory where your log files are located (you may have to change the file types in the Select Files and/or Folders dialog box) and select the file(s) to be analyzed. You can select multiple files and directories in the Select Files and/or Folders dialog box by using the Ctrl or the Shift keys. The file(s) and directories selected will be displayed in the Select the log file names for your site Input box. You can also add files from different directories by clicking the browse button again. Wildcards can also be used to specify files. You need to specify the file/s with the path (e.g. c:\logfiles*.log) and not just the path (e.g. c:\logfiles)

Use the defaults for File and Date format (Auto-detect), specify the profile type as Traffic and then click the Forward button.

Select to Show Report on screen and click the Forward button.

Click the Forward button on the next panel.

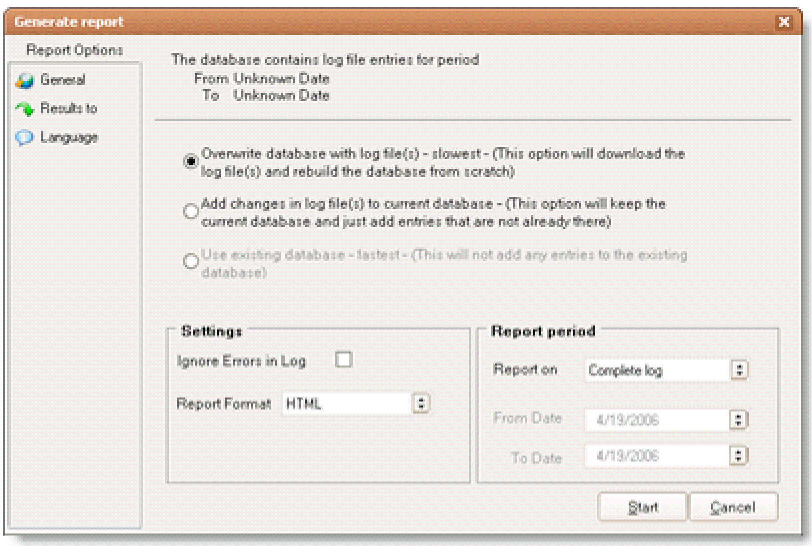
Click the Finish button on the next panel.

The site profile you have just created will be listed in the site profile list.

Please refer to the Working with Log Files section for step-by-step instructions to learn more on log file setup.

Populating the database

Select a site profile and click on the Report menu item (or double click the site profile). When you click the Report menu item, you are presented with three choices on how to populate the database:



1. Overwrite database with log file(s) - This option will download the log file(s), clear the report database for the profile and rebuild it from scratch using the data from the specified log files.

Duplicate date/time entries will be added to the database. This is the only option that you can use to analyze multiple sites and clustered servers.

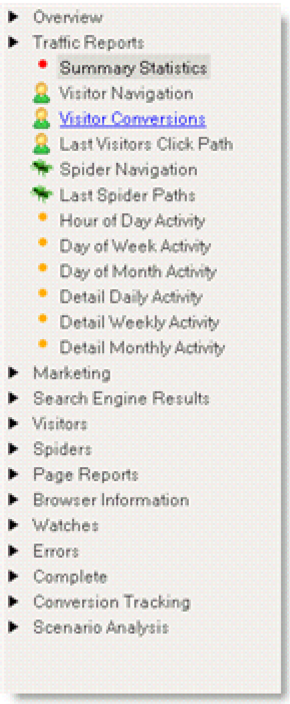
2. Add changes in log file(s) to the current database - This option will keep the current database and only adds entries that are not already in the database from the specified files. Duplicate date/time entries will not be added to the database. . For this to work, it must run through all specified log files to check for duplicate entries. It might appear that it is repopulating the database from log files already processed, but this is not the case. You will notice that it handles already processed log files faster than newly specified/changed log files. Please note that you should not use this option when analyzing multiple sites as one site. You cannot use the period specification when using this mode. Also note that if you change filters for a site profile using Fast mode, the database will be cleared first.

3. Use existing database - This will skip processing of log files and just generate reports from the existing database.

For now, select the first option and click on the Start button to start populate the database.

Generating reports

After the database has been populated, the report menu items (on the left) will be displayed. Reports (except for the Summary Statistics report) are not generated until you click on a menu item on the left.



You can also change the date selection by selecting a specific date or date range on the calendar visible at the top right.

Working with Log Files

What is a Log File

When a visitor's browser requests a file from a site, details of the request are recorded and stored in a log file on the web server. This log file may contain information like the date and time of the visit, the visitor's IP Address, the file requested, the status of the request, the size of the file requested, etc. It may also contain the referring site's information like the search engine (AltaVista, Google, Lycos, Yahoo, etc.), the search phrases or keywords used, etc.

Log files are text files and can be opened with a text editor. A typical log line (or hit) in a log file looks like this:

193.65.100.100 - - [25/Feb/2000:03:39:41 -0500] "GET /samples/clienterrors.htm HTTP/1.0" 200 3547 "http://www.altavista.com/cgi- bin/query?sc=on&hl=on&q=HTTP+error&kl=XX&pg=q" "Mozilla/4.0 (compatible; MSIE 5.5; Windows NT)"

Some web hosts do not supply log files with referrer information due to the extra disk space requirements. Many reports will not be available if your log files do not contain referrer information.

The above log line can be parsed into fields and contains the following information: 193.65.100.100 - This is the host name or IP address associated with the requester or visitor. Depending on the requestor's method of network connection, the host name or IP address may or may not identify a specific computer. Connections via many Internet Service Providers (ISPs) assign different IP addresses for each session, so the host name identifies only the ISP. If the computer is behind a proxy server or firewall, the proxy server or firewall IP address will be identified in most cases. The host name or IP address will identify a specific computer if that computer has a fixed IP address. By doing a Whois query or a Reverse Domain Name Server lookup, the domain name can sometimes be found from the IP Address. 25/Feb/2000:03:39:41 - Date/Time entry -0500 - Time Zone (of server). In this case 5 hours before GMT GET – Request type. In this case a Page Request /samples/clienterrors.htm - Filename Requested HTTP/1.0 - Protocol (and version) 200 - Status/Error code (200 = successful) 3547 - File Size (in bytes) http://www.altavista.com - Referrer /cgi-bin/query?sc=on&hl=on&q=HTTP+errors&kl=XX&pg=q - Search keywords or search phrases Mozilla/4.0 (compatible; MSIE 5.5; Windows NT) - Browser type, browser and operating system

Log files are usually stored on the server, possibly in a compressed format and may contain thousands of log lines. To analyze log files, you would have to:

Download the log files from your host's server. SurfStats Website Traffic Analyzer can also automatically download log files with the built-in FTP client.

Uncompress or extract the compressed log files.

Analyze the log files and create reports.

A log file analyzer like SurfStats Website Traffic Analyzer automates log file analysis for you and convert the log file entries to meaningful site activity reports.

Note that there are many different log file formats and no log analyzer can support all the different log file formats.

Where to get log files

If you are going to retrieve your log files from a remote server, you should contact your host or network administrator. You can either download the log files with a FTP client application or let SurfStats download the log files for you.

You would need the following information to set up a site profile that will tell SurfStats Website Traffic Analyzer where to get the log files and how to process them:

- How to log on to the host's server, i.e. the ftp address, login name and password.
- The folder where your log files is kept.
- The naming convention, whether it is Multi-file or combined log files, the archiving and rotation system used.
- What type of Operating System your host uses, e.g. Unix/Linux, Windows NT/2000/2003/XP, etc.
- The time zone in which the logs are recorded, i.e the GMT setting on the server.

Recommendations

We recommend:

- Use log files with referrer information.To get all the reports available in SurfStats Website Traffic Analyzer, your web server needs to log referrer information for SurfStats to analyze. Many hosts do not log referrer information due to the additional disk space it requires on their servers. Without referrer information, you cannot find out where your visitors came from and what keywords/phrases they used to find your site. We recommend that you request from your host to supply log files with referrer information (or find a host that does). See also the section on how to set up your server to generate log files with referrer (extended) information.

- Use combined log files if your web server is an Apache Servers. A log line from an Apache Combined log file looks like this (one long line):
  193.65.100.100 - - [25/Feb/2000:03:39:41 -0500] "GET /samples/clienterrors.htm HTTP/1.0" 200 3547 "http://www.altavista.com/cgi-bin/query?sc=on&hl=on&q=HTTP+errors&kl=XX&pg=q" "Mozilla/4.0 (compatible; MSIE 5.5; Windows NT)"

- You also get multi-file logs that contain hit information in separate log files, e.g. an access_log, a referrer_log and an error_log. This is an old format, takes up more space and is more difficult to manage when you want to analyze over longer periods. We recommend that you always use combined log files. Request from your host or network administrator to change the format.

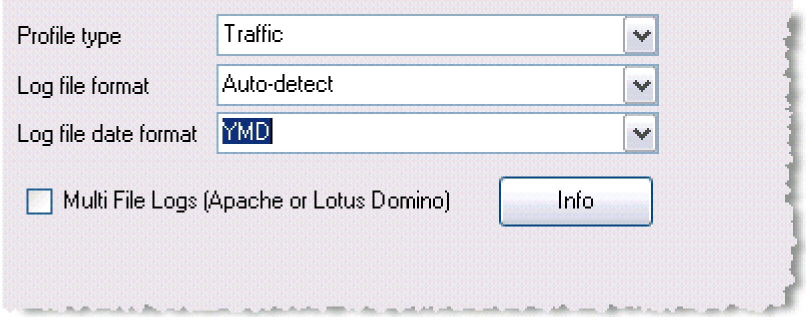
Server set-up

To get all the reports available in SurfStats Website Traffic Analyzer, you need to have referrer information in your log files. The following is a very basic description on how to set up extended logging to get referrer information in log files in Apache and Internet Information Server (IIS) servers:

Apache

Please refer to http://httpd.apache.org for detailed documentation on the Apache web server.

For Apache log files specify the **Log file date format** explicitly on the **Log Format** tab by editing the site profile.



Some site hosts supply their clients with separated log files, e.g.:

- access log - this contains the date/time and file accessed
- referrer log - this log file only contains the referrer information
- agent log - this log contains the browser or "agent" information.

With this configuration the web server logs to three separate files, instead of one combined file. If possible, request combined log files from your host as these files are easier to manage.

If you have control over the server you can configure Apache to generate combined log files by editing the "httpd.conf" file in the /Apache/conf/ directory and change entries as follows:
Specify the log file format with the LogFormat directive, example: LogFormat "%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i\""

Specify that you want a combined log file stored and the location, for example: CustomLog logs/combined.log combined.

In SurfStats you need to set up the profile to handle separate log files as follows:

- When specifying this log file type, only select the access log. The other log files will be added automatically.
- You should also enable the Multi File Logs... option on the site profile.

IIS

We suggest using the W3C Extended Log File Format as it is the only IIS5/6 format that logs Browser/Platform and Referrer info on IIS (Internet Information Server). SurfStats also supports the other available IIS types. To set up logging in IIS do the following:
Click on Start - Programs - Administrative Tools - Internet Service Manager - select your server and web site- right click on it and select Properties - Choose Active log format (W3C Extended Log File Format) - Click on Properties - Extended Properties and select the fields to log.

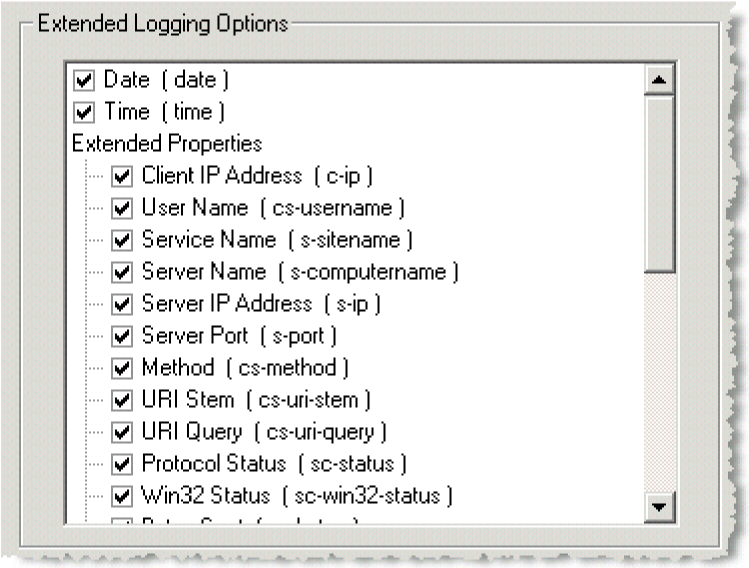
At least the following fields should be logged:

- Date
- Time
- Client IP address (c-ip)
- User Name (cs-username)
- Server IP address (s-ip)
- Server Port (s-port)
- Method (cs-method)
- URl Stem (cs-uri-stem)
- Url Query (cs-uri-query)
- Protocol Status (sc-status)
- Bytes sent (sc-bytes)*
- Bytes received (cs-bytes)*
- User Agent (cs(User-Agent))* - for browser information
- Cookie (cs(Cookie))* - for cookie information
- Referer (cs(Referer))* - for referrer information

* By default is NOT logged by IIS.

Of particular importance is the Protocol Status field. Sometimes logging of this field is turned off with the result that SurfStats produces no reports. This field contains the status code (for instance 200 for success, 404 for page not found etc.) for the HTTP request and is crucial in log file analysis.

We recommend the logging of **all** the fields on the **Extended Properties** tab. Stop and start the web server after editing the properties to log. Create an entry in the log file by accessing your web site with a browser and check if the entry is correctly logged in the log file.



Please note that IIS reserves and locks a 64 KB block in the current log file it is logging to. The current 64 KB block in use cannot be added to the reporting database. Some clients reported an error "Database empty" and expected this information to be added to the database.

Also note that the default setting for ASP buffering is ON in IIS5/6. This causes the sc-bytes field to be logged as 0 as the web page is served from the web server's cache. You should switch ASP buffering OFF to log this field. If you do not do this, you will not be able to track bandwidth in reports on ASP files.

Specifying log files

Date Macro Samples

Previous Top Next

You can use date macros to specify log files relative to the computer's system date. Let's assume that the current system date is June 30, 2012 and the naming convention of the log files on the server is exyymmdd.log, e.g. ex120630.log for a log file of June 30, 2012. It is also assumed that the log files are located in a directory called logs. You can then use, for example:

- %Date-1% to specify 1 day before the system date
- %yy% to specify the year of the calculated date (12)
- %mm% to specify the month of the calculated date (06)
- %dd% to specify the day of the calculated date (29 – assuming %Date-1%)

If you want to download the new log file for June 2012 every day, specify a file as /logs/ex1206%dd%.log (for instance). On the 30th day, the %dd% will be substituted by the current system day, namely 30 and the file specified will be /logs/ex120630.log.

If you want to specify all files for the current year and month, you can specify it as /logs/ex%yy%%mm%*.log. The file specification generated will then be /logs/ex1206*.log where the * will include all days of the month.

Assuming the same log file names as above If you want to specify log files of today, yesterday and the day before yesterday's log file (assuming a system date of June 30, 2012), you can specify the log files using three entries as follows:

- /logs/ex%yy%%mm%%dd%.log will yield /logs/ex120630.log (Today's log)
- /logs/ex%Date-1%%yy%%mm%%dd%.log will yield /logs/ex120629.log (yesterday's log)
- /logs/ex%Date-2%%yy%%mm%%dd%.log will yield /logs/120628.log (2 days ago)

Suppose your log file directory and format looks like this for example: logs/access.22Nov2012.gz and the current system date is 28 November 2012. To set up to download a log file of last Tuesday, specify it as follows:

- /logs/access.%Date-6%%dd%%Mon%%yyyy%.gz

In this example %Date-6% calculates the date relative to the system date (28 Nov 2012 - 6 days = 22 Nov 2012)
%yyyy% formats the calculated system date's year (2012)
%Mon% formats the calculated system date's month (Nov)
%dd% formats the calculated system date's day (22)

You have to specify every file or a selection of files that should be downloaded.
If you want to analyze a log file of two days ago and the log file looks something like this /logs/access-log.12-105.gz
where 12 is the day, and today is the 14th, you can specify it as

- /logs/access-log.%Date-2%%dd%-105.gz

Assuming today is October 26, 2012 and your log file's specification is exyymmdd.log. To specify all the log files for last month, the date macro should be resolved as ex1209*.log. Use the Date Macro Editor to experiment how the date macro works and how to get this specification. In this case the specification should be:

- /logs/ex%date-28%%yy%%mm%*.log

For %date-28% 28 days will be subtracted from today's date to evaluate to 120928.log. 28 was subtracted to ensure that last month or 09 is the result. If you process the log files on the 15th of each month then specify 16 as the number of days to subtract.

Using the File option

Using this method, you can specify log files located in:

- A folder on your local Computer.
- A shared folder on a Windows computer on the Local or Wide Area Network.
- Any Unix or Linux computer running NFS on a Local or Wide Area Network.
- A folder mapped with a third party mapping utility (for instance netdrive).
- A WebDAV folder allowing you to access folders over the HTTP protocol.

Select the site profile to set up, click on the Edit profile menu item, then on the Input tab, then on the Browse button. You can then browse to the drive and directory where the log files are located to select it or you can type the file and path name directly, e.g. c:\logs\site1\ex030101.log.

You can specify a mixture of uncompressed and compressed (.gz, .zip, etc.) files. You can use wildcards, e.g. access*.gz and you can use date macros to specify file names relative to the current system date.

You can specify compressed or uncompressed files and you can select multiple files from multiple drives and directories. Note that you cannot specify compressed files within compressed files, e.g. June2001.zip file which contains every day's compressed file. You can use a June2001.zip file with every day's uncompressed log file in it.

The following sample directory structure is suggested for easy file specification:

C:\Logs			
	\Site1		
		\Jan2001	
			\log1.zip \log2.zip \...log31.zip
		\Feb2001	
			\log1.zip etc.
	\Site2		

You can then easily specify the log files for analysis when you edit the site profile.

Note that you can edit the file source directly or you can use the Browse button to browse to a directory and select a single file or directory as well as multiple files or directories.

Please ensure that you have access rights to the locations where you wish to access the log files from.

Using the FTP option

FTP (File Transfer protocol) is used to upload or download files to web servers. SurfStats Website Traffic Analyzer has it's own built-in FTP client. With the FTP File Retrieval method, you can specify files to retrieve on:

- Your local Computer (By using a local FTP server on your computer).
- Any Windows or NT/2000/2003/XP Computer (All Versions) on a Local or Wide Area Network or over the Internet that is running a FTP Server.
- Any Unix, Linux , Windows, Windows NT or other computer that has a FTP Server on a Local Network, Wide Area Network or over the Internet

You will have to find out the following information from your network administrator, host or Internet Service Provider to set up FTP in SurfStats Website Traffic Analyzer:

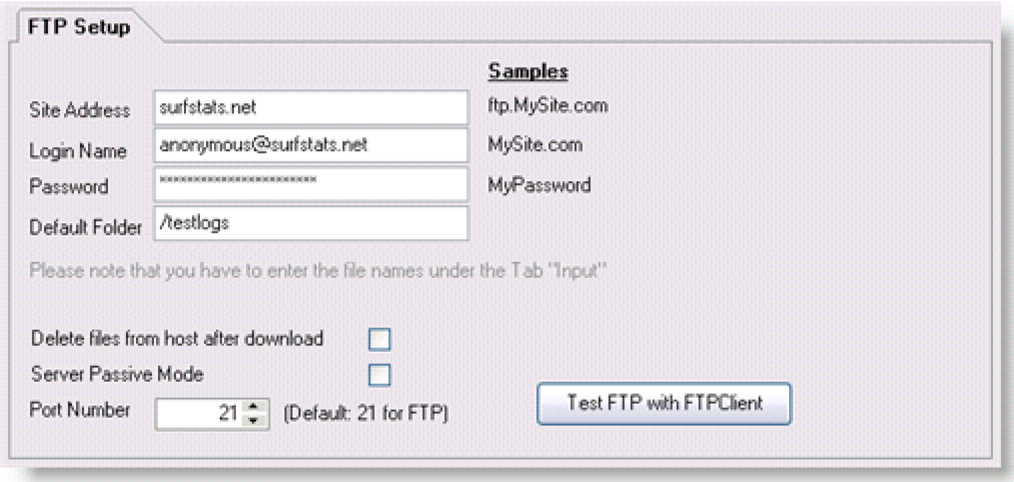
- The host name or IP address of the server
- The login name and password to access your log files
- The naming convention of your log files
- The rotation of your log files (when are they changed) as they may be renamed and compressed by the web server.

Note that the remote server must have a FTP server running.

Once you have the required information, you can then set up manual or automatic (unattended) retrieval of your log files and report generation.

Auto FTP

Select the site you want to set up for FTP. Click on the **Edit Site Profile** menu item, then on Site and set it to AutoFTP. Select the FTP menu item and set the Site Address, Login Name, Password and use the full path without the filename as the Default Folder (for instance /logfiles/)



Connect to the Internet (if you are not already connected). Click on the **Test FTP with FTPClient** button. Click **Connect** to connect to the FTP Server.

In the file list double-click a folder (if you are not already located in the correct folder). Click on a file to select it when the log files are displayed in the file list and right-click, followed by **Download**. The filename will be selected in the **Save As** dialog box. Click Ctrl-C to copy the filename (only the name). Disconnect from the server and close the Manual FTP client.

Now click on the Input tab where the absolute path and file name(s) of the log files should be entered. Click on the input entry panel and type (or paste) the filename. Edit the filename.

Some examples:

/mysite-logs/old/200012/*.gz will download all .gz files from the directory /mysite- logs/old/200012/

/mysite-logs/access-log will download only the access-log file from the directory /mysite-logs/.

/logfiles/W3SVC34/ex*.log will download all the log files with names starting with ex and having an extension of .log from the FTP directory /logfiles/W3SVC34.

You can add more complicated file specifications linked to your system's current date later with the date macro editor (Edit Site - Input).

Click on the **OK** button at the bottom of the dialog box.

Note: If the server is a Linux/Unix machine then multiple files should be specified as **access*** and not **access*.*** as in Windows and DOS. Some servers do not accept wildcards (e.g. *.*) as parameters. In this case, use the date macros to specify log files.

Note: Some server hosts create a file link (Unix terminology) to your log file(s). If there is a link use the exact path of where the log file(s) is/are located instead of the link, because SurfStats would not be able to interpret the link. If for instance your user name is JoeBloggs your home directory could be set up by the hosting company as /home/JoeBloggs. The server host could create a link in the /home/JoeBloggs directory called access-log that links to /www/logs/JoeBloggs-access-log which is the real physical file. Use /www/logs/ as the default folder on the FTP tab and /www/logs/JoeBloggs-access-log as the filename on the Input tab in this case.

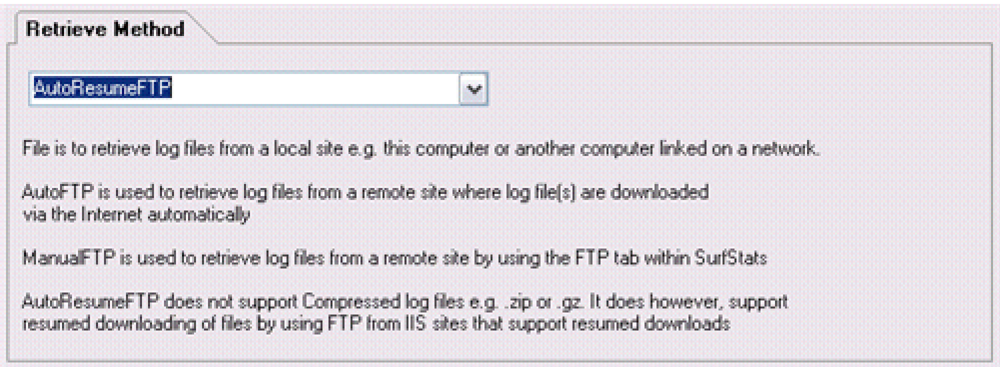
The downloaded log files are placed in the \Program Files\SurfStats<Version>\Sites\<ProfileName>\UnProcessed directory (where <ProfileName> refers to the name of the site profile). After the parsing process and updating of the database, the log files are moved to the \Program Files\SurfStats<Version>\Sites\<ProfileName>\LogFiles directory. You can specify this directory using the File option if you need to run an analysis on the log files again.

Please note that from V7 retrieval of log files with ODBC and HTTP is not supported anymore.

AutoResumeFTP

The AutoResumeFTP option is similar to the AutoFTP option. The only difference is that it can resume the download of files if it gets interrupted for some reason.

The method of retrieval can be specified by editing the particular profile and then selecting the **Site** menu item on the left. In the Retrieve Method panel select the method of retrieval from the dropdown box.



Note that in order for AutoResumeFTP to work the FTP server must support resumed downloads.

Selective file specification

You can use the * and ? wild card characters to selectively specify files as in the following examples:

c:\logs\site1\ex*.log
c:\logs\site1\ex9901*.log
c:\logs\site1\ex9901??.log
ftp.mydomain.com/logs/ex*.log

If you need to, you can specify multiple log files on one line separated by the semi-colon (;) character. Also refer to the Date Macros section on how to specify log files relative to the computer's system date. Please note that some servers do not allow you to use wildcards to specify log files and you would have to specify the log files explicitly or use date macros.

Date Macros

Date Macro Basics

Date macros can be used to selectively specify log files for analysis and also to create sub- directories for report output relative to the computer's current system date.

You can use date macros to selectively specify log files related to the current system date. The date, day, week, month or year macro can be used in strings and the program will then substitute the current system date when the specific string in the macro is encountered. Note: The %Date-X% macro only calculates the date and does not give any output.

The following date macros functions are available:

Macro	Specify As	Description
Date	%Date-X%	Today's date minus X days. This macro only calculates the date relative to the current system date. You need to specify the format of the day, month and year as well (using %dd% etc.).
Day	%dd%	The day of the month. 2 digit, e.g. 09
Week	%ww%	The week of the year, the week starts on a Monday and ends on a Sunday. 2 digit, e.g. 03. The %ww% macro is normally only used to specify a sub-directory for outputting reports.
Month	%mm%	The month of the year. 2 digit, e.g. 01 for Jan
	%mon%	Case sensitive, e.g. jan
	%Mon%	Case sensitive, e.g. Jan
Year	%yy%	2 digit, e.g. 12 for 2012
	%yyyy%	4 digit, e.g. 2012

Date Macro Samples

You can use date macros to specify log files relative to the computer's system date. Let's assume that the current system date is June 30, 2012 and the naming convention of the log files on the server is exyymmdd.log, e.g. ex120630.log for a log file of June 30, 2012. It is also assumed that the log files are located in a directory called logs. You can then use, for example:

- %Date-1% to specify 1 day before the system date
- %yy% to specify the year of the calculated date (12)
- %mm% to specify the month of the calculated date (06)
- %dd% to specify the day of the calculated date (29 – assuming %Date-1%)

If you want to download the new log file for June 2012 every day, specify a file as

/logs/ex1206%dd%.log (for instance). On the 30th day, the %dd% will be substituted by the current system day, namely 30 and the file specified will be /logs/ex120630.log.

If you want to specify all files for the current year and month, you can specify it as /logs/ex%yy%%mm%*.log. The file specification generated will then be /logs/ex1206*.log where the * will include all days of the month.

Assuming the same log file names as above If you want to specify log files of today, yesterday and the day before yesterday's log file (assuming a system date of June 30, 2012), you can specify the log files using three entries as follows:

- /logs/ex%yy%%mm%%dd%.log will yield /logs/ex120630.log (Today's log)
- /logs/ex%Date-1%%yy%%mm%%dd%.log will yield /logs/ex120629.log (yesterday's log)
- /logs/ex%Date-2%%yy%%mm%%dd%.log will yield /logs/120628.log (2 days ago)

Suppose your log file directory and format looks like this for example: logs/access.22Nov2012.gz and the current system date is 28 November 2012. To set up to download a log file of last Tuesday, specify it as follows:

- /logs/access.%Date-6%%dd%%Mon%%yyyy%.gz

In this example %Date-6% calculates the date relative to the system date (28 Nov 2012 - 6 days = 22 Nov 2012)
%yyyy% formats the calculated system date's year (2012)
%Mon% formats the calculated system date's month (Nov)
%dd% formats the calculated system date's day (22)

You have to specify every file or a selection of files that should be downloaded.
If you want to analyze a log file of two days ago and the log file looks something like this /logs/access-log.12-105.gz
where 12 is the day, and today is the 14th, you can specify it as

- /logs/access-log.%Date-2%%dd%-105.gz

Assuming today is October 26, 2012 and your log file's specification is exyymmdd.log. To specify all the log files for last month, the date macro should be resolved as ex1209*.log. Use the Date Macro Editor to experiment how the date macro works and how to get this specification. In this case the specification should be:

- /logs/ex%date-28%%yy%%mm%*.log

For %date-28% 28 days will be subtracted from today's date to evaluate to 120928.log. 28 was subtracted to ensure that last month or 09 is the result. If you process the log files on the 15th of each month then specify 16 as the number of days to subtract.

Date Macro Editor

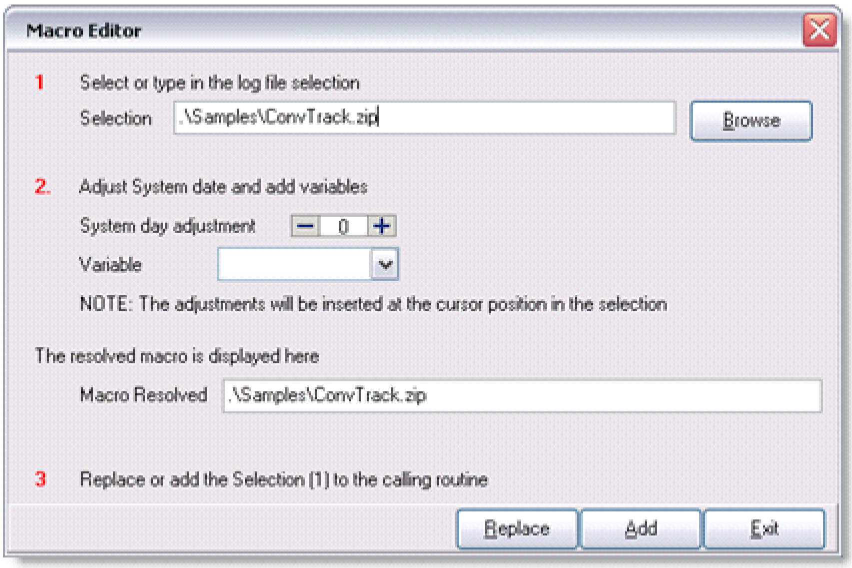
The Date Macro Editor is a dialog box used to assist in specifying date macros. Date macros gets transformed to file specifications when the SurfStats application runs to process log files and generate reports.

Please note that some servers do not accept wildcards (like *.* or *.???) when you try specify file names to download. In situations like these you would have to explicitly specify file masks that do not contain wildcard characters.

You access the Date Macro Editor by clicking on the **Macro Editor** button from the Input tab (where you specify log files to be analyzed) on the Edit profile screen.

To add a new log file specification or mask, make sure that no text is selected and that you are at the end of the list. If you want to replace an existing entry from the log file name list, select the entry first by double-clicking it.

Clicking the **Macro Editor** button brings up the Date Macro Editor as follows:



If you click the **Browse** button (next to step 1) you are presented with a Browse for Folder dialog that you can use to select the folder name part of the complete path (for example C:\WINNT\system32\LogFiles\W3SVC1\).

Clicking the "-" or "+" next to the **System day adjustment** field will offset the number of days in the file specification and add the text %date-System day adjustment value% to the **Selection** (next to step 1) field, at the current cursor position in the **Selection** field. So for instance if you specify –2 (by clicking "-" 2 times) for **System day adjustment** then %date- 2% will appear in the Selection field at the current cursor position. Note that the %date-2% will not be part of the final name specification it is only used to **calculate** a date relative to the current system date.

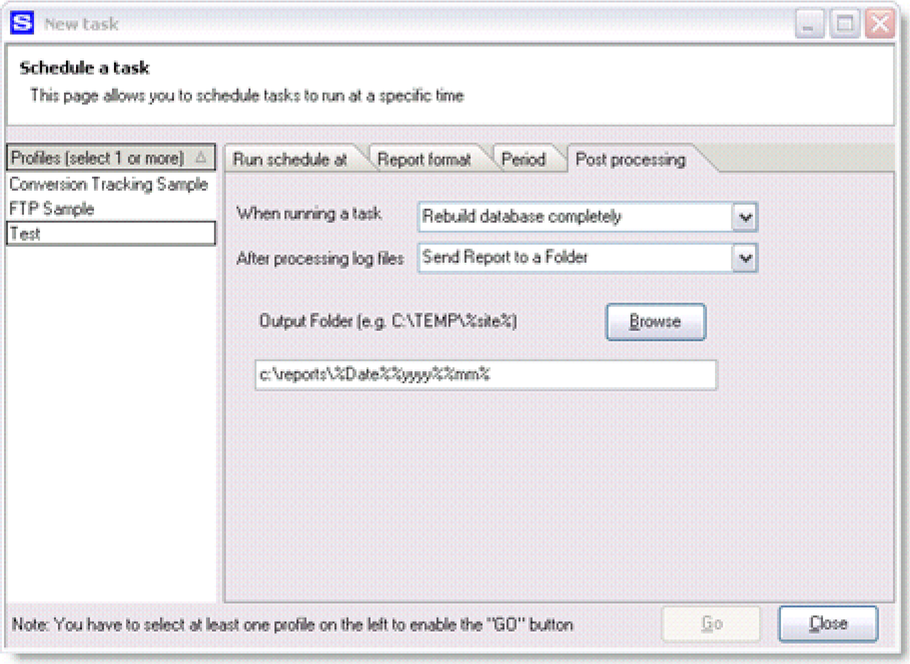
Selecting a variable from the **Variable** selection box will add the selected variable (for instance dd) at the current cursor position in the **Selection** field (next to step 1). A % character will be added before and after the variable. If you select (for instance) dd from the **Variable** selection box then %dd% will be added in the **Selection** field.

The resolved date macro, i.e the file name that will be used as file specification by SurfStats, is displayed in the **Macro Resolved** field as changes are made to steps 1 and 2.

After you are satisfied with the changes done you can click Replace to replace the selected text (or insert the text at the current cursor position if no text selected) on the Input tab (of the site profile edit screen) or Add to add the specification at the end of the list on the Input tab.

Date Macros and Report Output

Similar to using a date macro in specifying log files a date macro can be used to generate a folder name for report output from the SurfStats Scheduler.When selecting to output reports to a folder a date macro can be specified to generate an output folder name.



For instance if today's date is Nov 8, 2005 a date macro of c:\reports\%Date%%yyyy%%mm% will evaluate to c:\reports\200511 and generate reports to a folder with that name.

Please note that the parent folder (for instance c:\reports) must exist.

Advertising analysis

SurfStats Website Traffic Analyzer can help you to track the number of ads that was displayed (ad views) and the number of visitors that clicked on particular ads (ad clicks) on web pages. This advertising traffic statistics can be helpful if you are selling banner advertising space on your site.

Hits from external links

A banner ad is usually an image with a link to the content located at the advertiser. When a visitor clicks on the banner ad, the visitor is then redirected to the advertiser's web page that is usually an external page or file (files not on your web server). Requests for external files do not generate hits in your server log files and can therefore not be normally tracked. To generate hits in log files to measure clicks to external pages, you would have to create a redirect page that is located on your server and change the original link to this redirecting file. A click on the banner ad (or external link) will then request the redirecting page on your server, the server will log this entry in the log file and the user will be directed to the advertiser’s web site (by code on the redirect page). This happens so fast that the user does not realize that a redirection has occurred.

You can then find out the exit pages from your site or clicks on ad banners by measuring the number of times that this re-directing file was requested. If for instance you want to place an ad for SurfStats on your web page and want to see how many users click on this ad, you can do the following:

Use an Html refresh page for every ad

An easy way is to create a new page for the ad or external link, call it, e.g. surfstats.htm and place the following html code in it:

```
<html>
<head>
<meta http-equiv="refresh" content="0; url=http://www.surfstats.com/thecontent.asp">
</head >
</html>
```

You then link (with a hyperlink) the advertisement on your web page to this re-directing page (surfstats.htm) on your site. When the page with the above HTML code gets called the meta tag will cause a redirection to the page thecontent.asp that contains the content for the ad. You would have to create a page for every ad that you want to track with this method.

When a visitor clicks on the ad, a hit to the surfstats.htm page will get logged in the web server's log file and the visitor will then be re-directed to www.surfstats.com. To count the number of users that clicked on the banner, just count the number of hits for the re-direct page.

Use an ASP redirect

Active Server Pages (ASP) only works on Microsoft IIS web servers. An URL query can be used for re-direction by specifying the ASP file name, followed by a question mark, an equal sign and the destination file as in the following example:
adclick.asp?Ad=http://www.advertizer1.com/product1.htm

The adclick.asp page would then contain server-side redirect code to redirect the browser to http://www.advertizer1.com/product1.htm. For instance:
<%Response.Redirect Request.QueryString("Ad")%>

You can then set up **Separate parameterized queries to same page** on the site profile to track access to the re-directed pages.

Use the free re-direct program from AnalogX

This application works only on Windows servers. It is available from this link:
http://www.analogx.com/contents/download/network/cgirdir.htm

Ad Reporting

There are two reports to track advertising on web pages, namely:

Ad Views

This is the number of times that a banner ad (or image) was requested by the user's browser, i.e. the number of times that the ad was displayed (viewed) on the visitor's browser. To generate this report requires that you define a unique part of the graphic file name in the **Filename** field (Options - Advertisements) and that a hit to this file must be recorded in the log file. This can be a gif, png, jpg, cgi, and asp, etc. file.

Ad Clicks

This tells you how many times visitors clicked on the banner ad. To generate this report requires that a hit to the resulting page (or redirect page) must be present in the log file and the resulting page or a unique part of it must be defined in the Click String (Options - Advertisements).

Please note - Ad analysis will only work if you have set up the ads to track correctly and if you have referrer log files with hits to the defined ad files in it.

Tracking ads set up

You first need to set up advertisement tracking . You do this by selecting **Options - Advertisements** from the Site profile menu panel.

To add new advertisements click the **Allow edit** checkbox. This will allow editing in the table (or grid) listing the advertisements. To insert a new advertisement click on the table to give it focus and press the Insert key on the keyboard (or click the **Add** button). A new row will get created. Fill in the

Description, **Filename** and **ClickString** fields for the row. Note that the new row will get saved when you move to another row. To delete an advertisement click on the row and then click the Delete button. Click Accept when you're done.

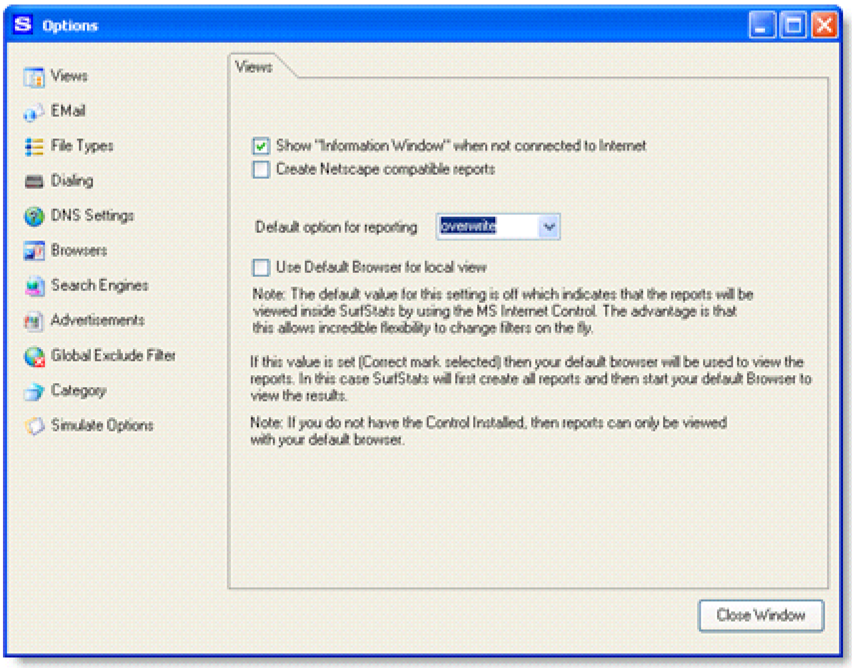
The **Filename** field is the name of the image (or advertisement) file on the page. The **ClickString** field is the file that will handle the redirection on your web site for the specific advertisement.

Please note that values filled in in the **Filename** and **ClickString** fields are case-sensitive. For instance if you specify surfstats.htm for the **ClickString** field and the real filename on disk is called SurfStats.htm, then SurfStats won't see it as an ad click.

Also note that you do not need to insert the complete filename or clickstring in the fields but only a unique part of it. It might be helpful to view the log file with the File Viewer, find the Ad Identification and then cut and paste this text string into the two fields.

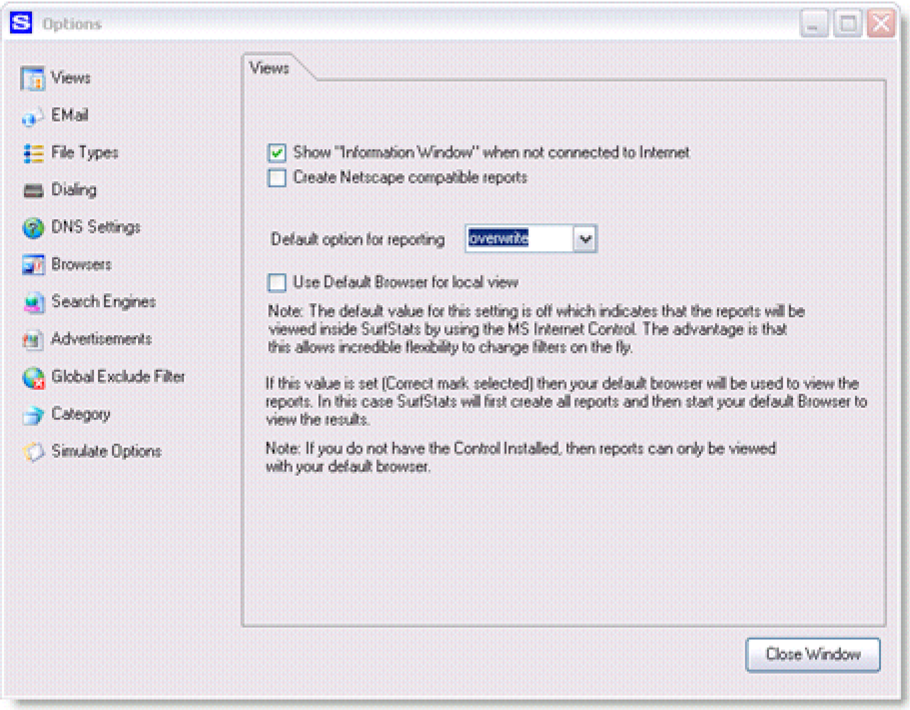
Changing global settings

You can set the following global report settings by clicking **Options** on the main toolbar when in Site profile mode.



Views

Sets up general view options.



Show Information Window when not connected to the Internet. This shows the Information Window when not connected to the Internet. If you have a permanent connection, clear this box.

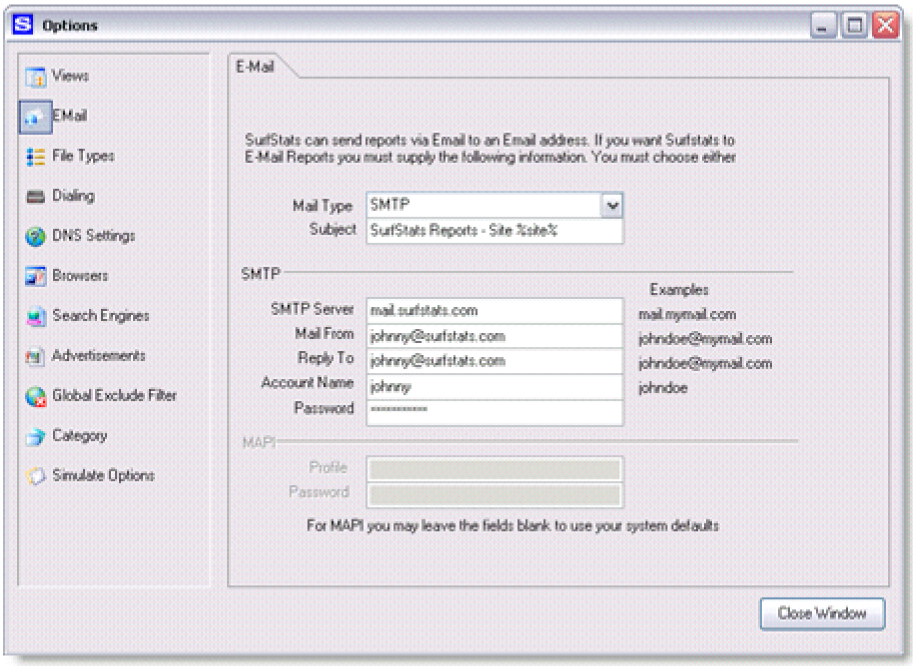
Create Netscape compatible reports. Netscape cannot display the shading used on graphs and reports. If this is selected, no shading will be generated on reports.

Default option for reporting. Specify which option is selected (**Overwrite Database**... or **Add Changes**…) when a site profile is double-clicked to generate reports.

Use Default Browser for local view. If selected, the default browser will be invoked to view reports. We suggest you use the latest version of Internet Explorer.

E-mail

This sets up global options for e-mailing reports.



Mail Type – SMTP or MAPI. Determines the mail protocol that SurfStats will use.

Subject – The subject of the e-mail containing the reports. Useful is the %site% macro that will get replaced with the profile name when reports are generated.

SMTP Server – The host name of the SMTP mail server, e.g. smtp.myisp.com.

Mail From – The From field in the e-mail message containing the reports.

Reply To – When the recipient of the e-mail containing the report replies this address will be displayed in the To field of the reply message, should the recipient of the email message reply.

Account name – Mail account on the SMTP server.

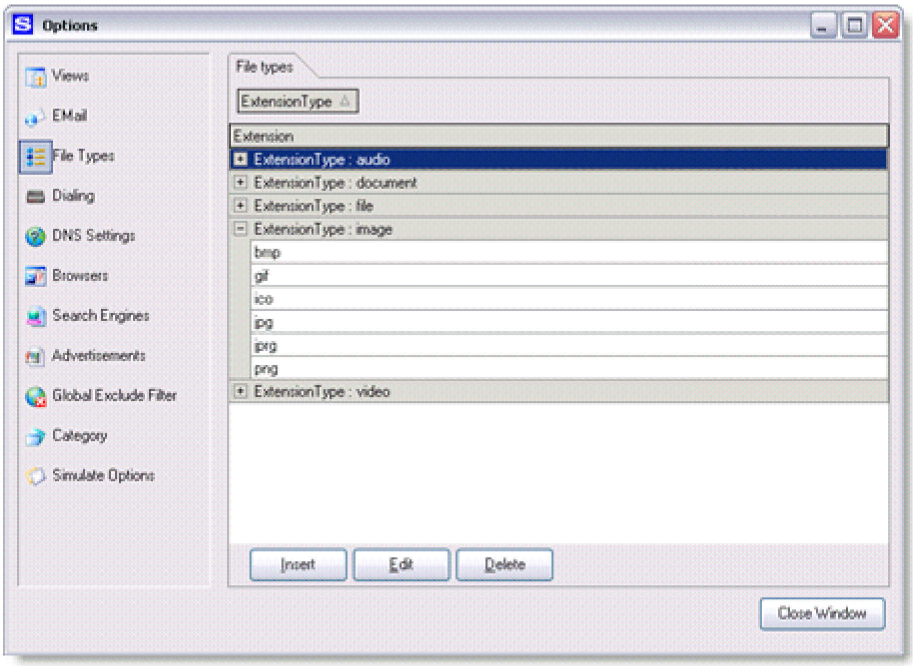
Password – The password to access the SMTP server.

Profile – The MAPI profile name.

Password – The password to access the MAPI profile.

File Types

File Types options are used to specify how SurfStats should classify a hit line when a specific file extension is picked up in a log line. For instance a hit line with .htm as extension is classified as a document (web page), one with a .pdf extension as a file, one with a .gif extension as an image etc. These options are used when the reports under the **Pages** category are generated. If a hit line is picked up in a log file and SurfStats cannot find the file extension in any of the extension types listed in **File Types** then the hit line is seen as a request to an **unknown** file.



The most common file and page types are tracked by SurfStats. You can add, delete and select file and page types to track. To add a file type click on the **Insert** button and select an **Extension Type** and specify a **File Type**. Do not use a full stop or dot (.) in front of the file suffix. To edit/delete a file type, select it and click on the **Edit/Delete** button.

Dialing

Specify whether the connection to the Internet is permanent or a dial-up connection.

If you connect to the Internet with a dial-up connection you can specify if SurfStats should initiate dialing or not. Note that this only works for connections classified by Windows as dial- up connections, not LAN connections.

DNS Settings

Specify the configuration for resolving IP addresses to DNS names.

DNS Lookup - You can set the number of threads that SurfStats should create when performing Reverse DNS lookups. Best is to stick with the default value of 100. Specifying a value of 100 for instance means that SurfStats Website Traffic Analyzer will attempt to do 100 reverse DNS lookups simultaneously.

DNS Ranges – Ranges of IP addresses that can be pre-defined and won't be looked up on the defined DNS Server when SurfStats performs a reverse DNS lookups. You can edit this table to add your own ranges (for example the IP address ranges on your Intranet). You can also export and import an IP address range table. The advantage of having DNS Ranges is that it speeds up the process of Reverse DNS lookups. **Please note** that you cannot use wildcards (? and *) when specifying DNS ranges.

From V7.0 of SurfStats Website Traffic Analyzer a separate **Visitor Information Database** is available with pre-resolved IP addresses based on Whois queries. After a new installation there will be no entries listed on the **DNS Ranges** tab, as the pre-resolved entries are present in the Visitor Origin

Database (if the Visitor Origin Database is installed and registered). The entries in the Visitor Origin Database cannot be seen or edited from SurfStats Website Traffic Analyzer, they are only displayed on reports.

Tip. The best way to get the resolved IP Address names is to set your server (or ask your host) to set up reserve IP lookup for your log files. The resolved names will then be placed in the log files and can then be used with Surfstats Website Traffic Analyzer.

Browsers

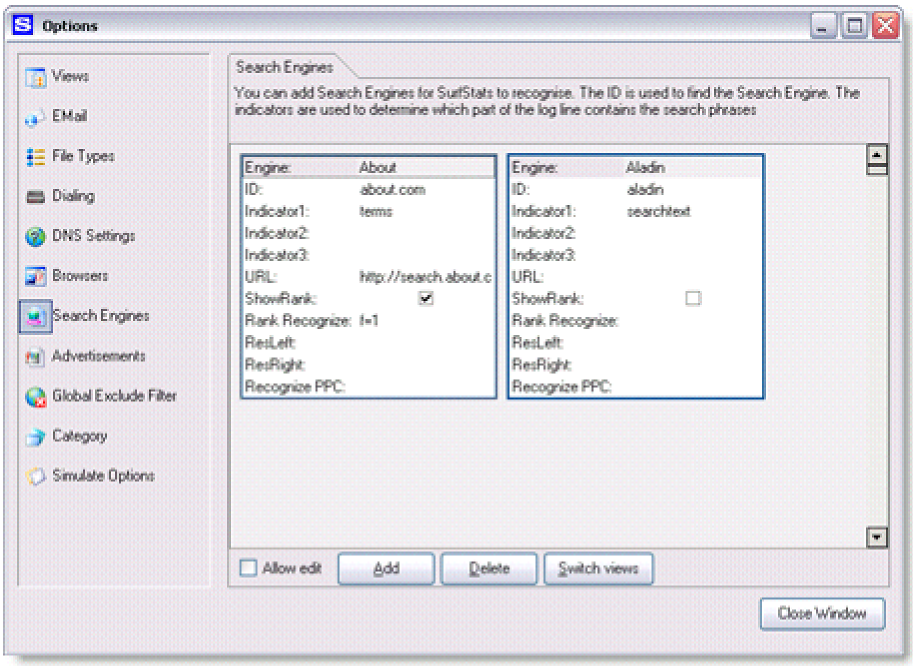
You can add, edit and delete footprint information that browsers, spiders and download managers will leave in log files, to be used when SurfStats analyzes log files when generating reports.

Spiders (sent out by Search Engines to crawl and index the web) are associated with a specific search engine and therefore has the **SearchEngine** field filled in, so when the browser footprint of the spider is picked up in the log line it can be associated with a specific search engine.

It is useful to view a relevant log file in the file viewer while modifying these entries. To add a Browser, click on the **Allow edit** check box, click on the list of Browsers and press Insert on the keyboard (or click the **Add** button). Add the necessary field descriptions for the row, move the cursor out of the row to save changes. Click **Close Window** when you're done.

Search Engines

SurfStats Website Traffic Analyzer can determine referrers from most of the major international search engines. Most countries have specific search engines and if you want to track referrers from them, you can add the Search Engine with the appropriate parameters to the list of Search Engines. Click on the **Options** menu item and then on **Search Engines**.



You can add, edit and delete search engines in order for SurfStats Website Traffic Analyzer to recognize it as a search engine. It is useful to view your current log file in the file viewer while modifying these entries. To add a Search Engine, click on the **Allow edit** check box, click on the list of Search Engines and press Insert on the keyboard (or click the **Add** button). Add the necessary field descriptions for the row, move the cursor out of the row to save changes. Click **Close Window** when you're done. We suggest that you browse with a browser to the search engine (for instance to www.google.com) that you want to add and enter a search phrase and click the search engine's **Search** button. From the URL generated in the browser's address bar you can then add the engine, its ID and indicators. Also refer to the existing search engine entries for examples. The field descriptions are:

Engine – This is the search engine name under which all the different variants of a specific search engine will be grouped.

ID – This is a variant of a specific search engine, e.g. "altavistacanada" which is the Canadian AltaVista search engine.

Indicator1 – Indicator3 - This is the footprint for the specific search engine variant to determine the search phrase used. You can determine this indicator by searching for a search engine referral and inspecting the search string. The indicator is usually one or more characters, preceded by "&" and followed by "=", e.g. in the following string:

http://search.msn.com/
results.asp/?co=200.1.4.10.&FORM=MSNH&RS=CHECKED&q=+log+file&v=1

there are three search elements namely FORM=, RS= and q=. In this case only the q should be used as a footprint. The keywords following the q= will then be used, namely log and file.

ShowRank - If this is checked the seach engine will be listed when the SurfStats Page Rank Tool is run.

The fields **Rank Recognize**, **ResLeft**, **ResRight** and **Recognize PPC** are also used by the SurfStats Page Rank Tool.

Tip: You can maximize the Search Engine dialog to display more search engine options at once.

Advertisements

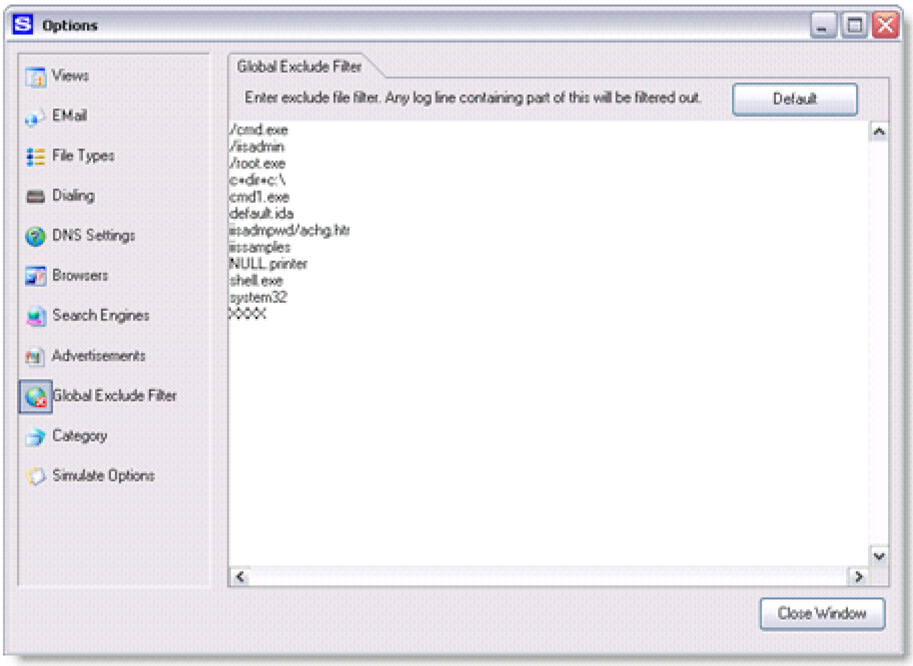
Please refer to the section on **Advertising analysis** on how to set up and track advertisements.

Category

Please refer to the section on Proxy Analysis for more information.

Global Exclude Filter

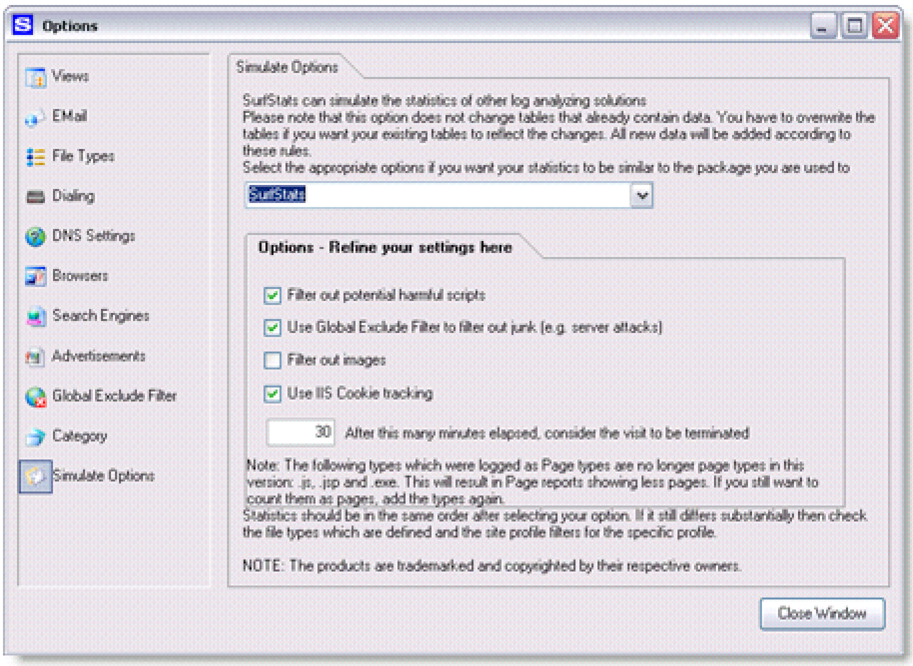
To access this option click **Options – Global Exclude Filter**.



If SurfStats processes a log file and encounter text in the line listed in the **Global Exclude Filter** it will discard the log line. This is useful for removing virus and worm entries. You can specify entries for the **Global Exclude Filter** editing directly in the list of filters. Click **Close Window** when you are done.

Simulate Options

To access this option click **Options – Simulate Options**. Options specified here control how SurfStats generates reports. Choose a package to simulate at the top if you want report generation to be similar to the package selected.



The following options can also be specified:

Filter out potential harmful scripts. SurfStats generates reports and show them in a browser, which could allow a malicious user to place <SCRIPT>, <OBJECT> etc. tags in hit lines in your log files causing the scripts or objects to execute when you view the reports.

Use Global Exclude Filter…Checking this option will determine if SurfStats will apply the **Global**

Exclude Filter when processing log files. If checked and SurfStats encounters a log line that contains text specified in the **Global Exclude Filter** it will discard the log line.

Filter out images. Checking this will make SurfStats discard hits to image files (as defined in **Options – File Types – ExtensionType: image**). Filtering out of image can dramatically increase report generation time. In general hits to image files do not contribute much to log analysis. This settings is the global setting and can be overridden per profile.

Use IIS Cookie tracking. Check this option to let SurfStats group log file hit lines with the same cookie information (even though they may be coming from different IP addresses) together. This is useful for visitors from AOL, which uses multiple IP Addresses for the same visitor, causing the visitor count to be inflated (not deliberately). To enable this feature cookie information must be logged in the log file. See the section on **Log files – Server Setup** for more on which fields to log. This only works for log files from IIS web servers.

Changing Appearance

Skins

V6.7 and some prior versions of SurfStats Log Analyzer made use of skins to change the appearance of the SurfStats Log Analyzer application. Skins were removed fromV7 of SurfStats Log Analyzer. The appearance of SurfStats Log Analyzer now solely depends on the Windows Scheme selected. In Windows 2000 this can be changed by right-clicking on the Windows Desktop, selecting **Properties** and changing the **Scheme** on the **Appearance** tab.

Report Appearance

Themes

You can change the backgrounds, fonts, colors and images on reports by applying a theme to it. You can also modify the existing themes or add your own design. Note that a theme is associated with a site profile, which means that the reports generated by SurfStats Website Traffic Analyzer for different site profiles can have different appearances.

Selecting a theme

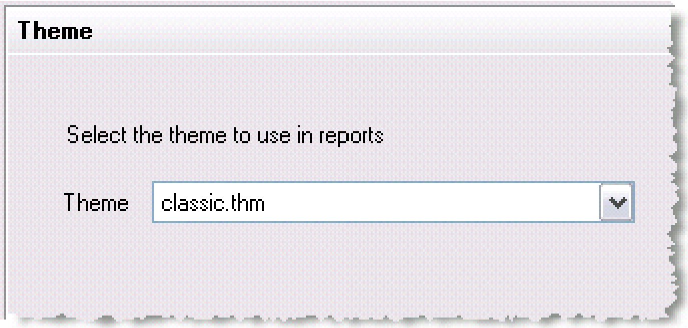
Click **Select Theme** on the top toolbar when in the report mode (a report is displayed in the viewer panel in this mode) to choose a theme from the list of themes displayed. After the Select Theme button has been clicked you can use the Up and Down arrow keys to quickly run through the themes and select the one you like. When you close the Themes window, the currently selected theme is saved as the theme for the site profile.

Modifying themes

The **Themes Editor** is used to modify or add new themes. The Themes Editor can be run by clicking the **Themes Editor** menu item under Tools on the left or by selecting **Themes Editor** from the **Tools** menu at the top.

You can then make changes to the elements of the theme. After you are done with the changes to the theme, click the **Save as** button, supply a name and exit the Themes Editor. You still have to select the new theme from the list of themes to make it the active theme of the specific site profile.

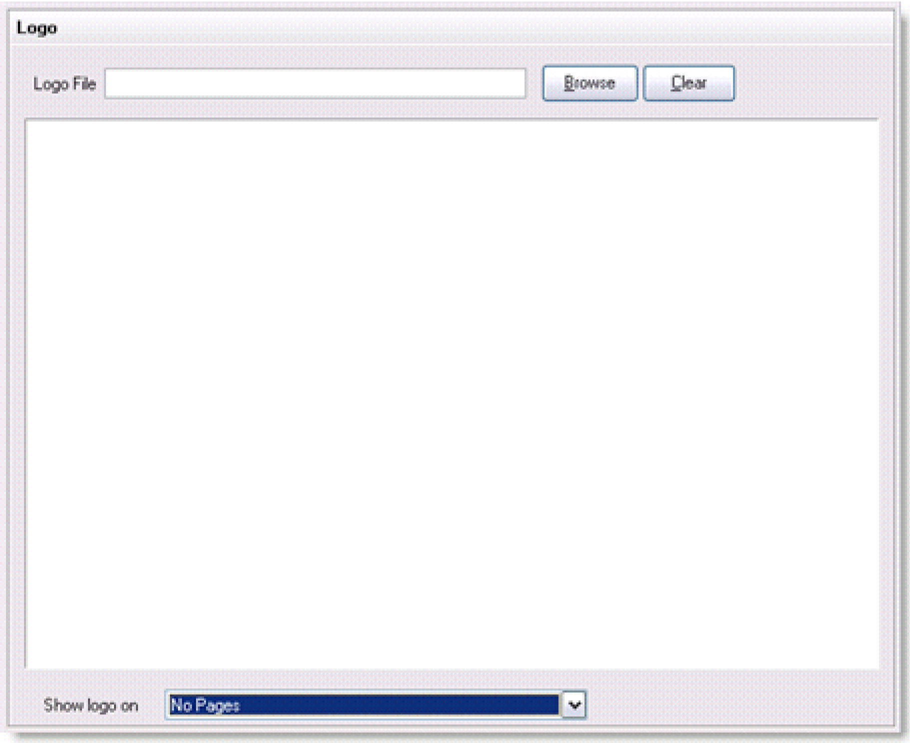
You can also change the theme for a specific site profile (before report generation) by selecting the profile, clicking on **Edit Site**, then on the **Theme** menu item.



Logo on reports
You can change the logo for every site profile that will be used on reports. To do so, select the profile to edit, then click **Edit Site Profile** and click the **Appearance** menu item. Click the **Logo** tab.
You can select a logo (your corporate logo for instance) and display it on all pages, only the first page or no pages at all. The headers and footers on reports cannot be changed from within SurfStats Log Analyzer. Please remember to select **Show logo** on at the bottom after the logo file has been selected.

Report Logos

You can change the logo for every site profile that will be used on reports. To do so, select the profile to edit, then click **Edit Site** and click the **Logo** menu item on the left.



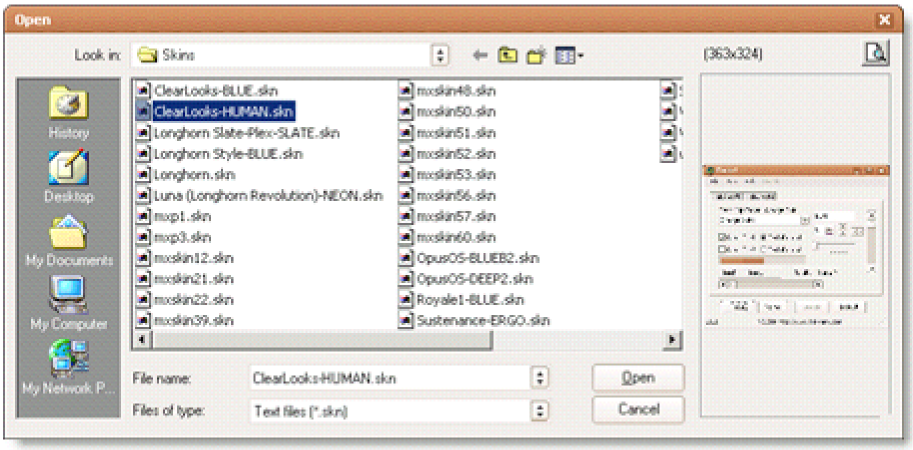
You can select a logo (your corporate logo for instance) and display it on all pages, only the first page or no pages at all. The headers and footers on reports cannot be changed from within SurfStats Website Traffic Analyzer. Please remember to select **Show logo on** at the bottom after the logo file has been selected.

Application Skins

Application skins can be used to change the appearance of the SurfStats Website Traffic Analyzer application.

From the **View** menu select the **Skin/Unskin** option to add skins or totally remove skins from the SurfStats Website Traffic Analyzer application.

Click the **Select a Skin** menu item to select the skin to use for the application. The skin selection dialog box appears. Select a skin by clicking the skin file name and clicking the **Open** button.



The selected skin will then be applied to the application.

Filters and Watches

Filters allow selection of certain hit lines to be excluded or included when processing log files to generate reports.

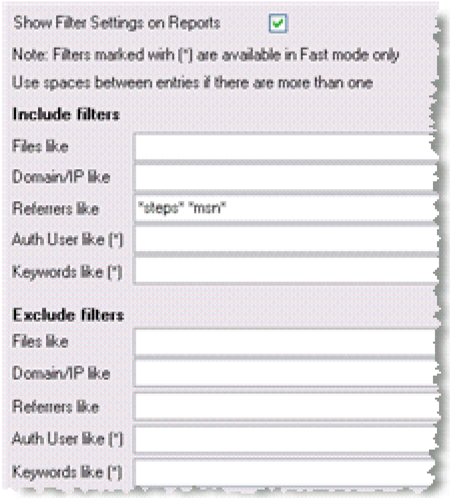
Client and Server Filters

About Client and Server Filters

Client or server filters can be specified per site profile. Client filters gets applied to characteristics of the visitor accessing the web site whereas Server Filters gets applied to characteristics of the web server hosting the web site. Filter settings can be accessed by editing the profile and selecting General Filters for client filters or Server Filters for server filters from the menu bar on the left.

Client Filters

If the General Filters option is selected the filter selection screen containing a set of include and exclude filters is displayed as shown below:



Server Filters
If the Server Filters option is selected the filter selection screen containing a set of include and exclude filters is displayed as shown below:

Including and Excluding

Filters can be used to either include or exclude log lines when SurfStats builds the reporting database used in reporting.

Include Filters

If a filter is specified under the Include Filters section in General Filters or Server Filters log lines matching the filter criteria will be added to the reporting database when the log files for the profile are processed.

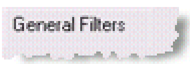
Exclude Filters

If a filter is specified under Exclude Filters section in General Filters or Server Filters log lines matching the filter criteria will **not** be added to the reporting database when the log files for the profile are processed.

Note: If a log lines meets the criteria of an include and exclude filter the log line will be excluded.

Specifying Client Filters

Client filters include or exclude hit lines based on attributes provided by the visitor's browser to the web server.



If a profile is edited and General Filters is selected from the menu bar on the left the following types of client include and exclude filters can be specified:

Files

A files like filter is used to include or exclude log entries based on the filename requested by a visitor to the web site. Some samples:

Filter	Will include or exclude
robots.txt	All requests for robots.txt the file requested by Search Engine spiders.
.js	All JavaScript files with a js extension, for instance utilities.js.
.gif *.jpg*	All image files with a gif or jpg extension, for instance okbutton.gif,
/john/	All files and folders located in the web directory john on the website.

As of V7.0 filters are case insensitive so for example an exclude filter that specifies ***/demo*** will exclude /Demo/myfile.htm and /demo/myfile.htm.

Domain/IP

An IP address identifies a visitor on the Internet. In most cases IP addresses are assigned automatically when the visitor connects to an ISP. Visitors can be filtered on based on the IP address they used to access a site. Some samples:

Filter	Will include or exclude
196.1.1.1	All requests from the visitor with the IP address 196.1.1.1.
209.*	All requests from visitors with an IP address starting with 209.

You can exclude your own visits to your web site if you know your IP Address.

Referrers

Use this type of filter to include or exclude log lines based on referrer, which is the site that directed the visitor to your website. For instance specifying *google* will include or exclude all log lines referred by Google. Referrer info needs to be logged in order for a referrer filter to work.

Authenticated Users

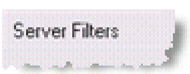
When a web server requests from a visitor to log into a website or into a protected area on a website the user name specified by the visitor is logged as one of the fields in the web server's log file. The Authenticated User's filter includes or excludes log entries based on this user name. Specifying *johnny* for instance as an include authentication user filter will only add log lines to the reporting database and resulting reports from the visitor that logged in with the user name johnny. The authenticated user name needs to be logged in order for an authenticated user filter to work.

Keywords

Keywords are the words typed in at search engines or other sites used to reach your website. The keywords used at the search engine is passed as part of the referrer info logged in the log file. For instance an include keyword filter of *analyzer* will only add log entries containing the keyword analyzer to the profile's reporting database. Referrer info needs to be logged in order for a keyword filter to work.

Specifying Server Filters

Server Filters can be set up to report on attributes of the web server, which is independent of the visitor's browser. This feature only works for log files from IIS web servers. Edit a profile and click the Server Filters menu item to set server filters.



IP

This filter allows you to specify the server IP address to filter on, which is the IP address of the virtual web server that served the request. For instance *196.1.1.1*

Port

The TCP port of the virtual web server that served the request can be filtered on. The default port that a web server serves requests on is 80. Specifying *20000* for example will only include requests received on port 20000.

Domain

This filter can be used to specify the domain that was requested. This filter is useful if multiple domains are hosted on the same virtual web server (on the same IP address and port). Specifying *www.john.com* for example will only include requests received for www.john.com.

Important Filter Tips

- Keep filters as short and as specific as possible as it slows down processing of log files.
- Ensure that the relevant fields are logged for the type of filter you are using. Please refer to the Client Filters and Server Filters section for more information.
- To test if a specific filter works first just specify the filter as an include filter and then specify it as an exclude filter.
- After changing filters use the "Overwrite Database..." method of report generation.

Watches

Watches are used to track access to certain files, IP addresses and referrers. For example if you sell three products online you can have a watch for each of the three product pages. Watches were first added in the SurfStatsLive product, as a live statistics product, in most cases has no way for a report end user to set filters.

Essentially a watch works the same way as a client filter with the following differences:
- A watch belongs to a watch group and can have a descriptive name, which is not the case with a filter. SurfStats Website Traffic Analyzer allows you to specify nine watches with three watches per watch group.
- A watch has an associated report whereas a filter is applied to all log lines processed and affects all the reports for the profile.

You can edit the group name of a watch group by clicking twice (slower than double-clicking) on a group name and editing it.
You can edit the watch name by clicking the watch name two times and editing it.
The * character needs to be specified before and after the watch filters, for instance *CompanyChosen.aspx* as a Files like filter.

Please refer to the Client Filters section for more information on specifying client filters for a watch.

After report generation the reports under the Watches report category display the watches reports.

Content Management Systems

SurfStats and CMS

SurfStats can replace page names and page parameters which makes the reports generated by the analysis of log files from web servers with Content Management Systems (CMS) easier to understand and to interpret. For instance:

Product Information Page:Category=Color Laser Printers&Product=Minolta ML309

is easier to understand than for instance:

productpage.aspx?prodcat=X267&prodid=M193759

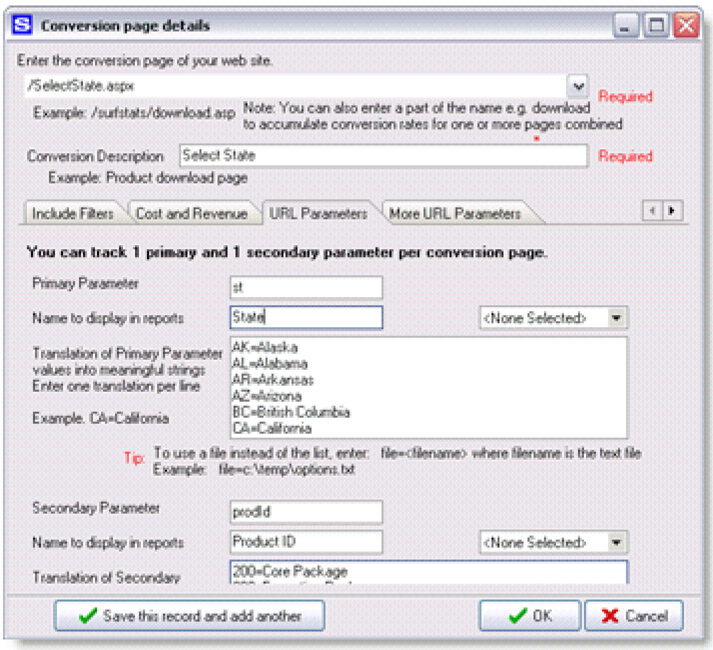
Pages and parameters for CMS are configured in Conversion Tracking. Please refer to the Configuring SurfStats section for more information.

SurfStats cannot connect directly with the database or use the API of a CMS system. However you can periodically generate export files from your CMS system and SurfStats will import these files and perform replacements whenever reports are generated for the profile. This approach enhances flexibility and allows SurfStats to work with virtually any CMS system that can generate an export file of the content. Please refer to Exporting Data from your CMS for more information.

Configuring SurfStats for CMS

Pages and Page Parameters for CMS are configured in the Conversion Tracking section for the site profile. Select the profile and click the Conversion Tracking menu item on the left to display the Conversion Tracking screen on the right showing a list of conversion pages defined.

Then click the **Add** button at the bottom, which will display a screen similar to the following, allowing you to add a conversion page:



At the top for the **...conversion page...** select or type the name of the page name you want SurfStats to recognize and replace. You can only select from the drop down list if reports were previously generated for the specific profile.

In the **Conversion Description** field specify the text that you want SurfStats to display in reports for this page. SurfStats will substitute this text for the the page name specified in the **...conversion page...** field.

To include only log lines when a page parameter has a specific value (for instance st=CA) specify the parameter(s) and value(s) on the **Include Filters** tab. The **Cost and Revenue** tab is not applicable to CMS.

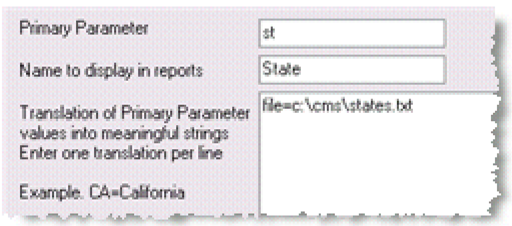
On the **URL Parameters** and **More URL Parameters** tab you can specify the parameter names and their replacement values. For example - in the sample screenshot above:

- If the page name *SelectState.aspx* is found it will be replaced with the text *Select State*.
- If a parameter called *st=CA* is found (for the page *SelectState.aspx*) it will be replaced with the text *State=California*. So *st* will be replaced with *State* and *CA* will be replaced with *California*.
- If a parameter called *prodid=200* is found (for the page *SelectState.aspx*) it will be replaced with the text *Product ID=Core Package*. So *prodid* will be replaced with *Product ID* and *200* will be replaced with *Core Package*.

In the **Translation of Primary Parameter values...** field you can specify a file name that contains the replacement values for the parameter. Please refer to Exporting Data from your CMS for more information.

Exporting Data from your CMS

Instead of specifying the parameter and replacement values on Edit Profile - Conversion Tracking - URL Parameters tab you can specify a filename that contains the replacement values for the specific parameter using the syntax file=<yourfilename>, for instance file=c:\cms\states.txt.



At the time that SurfStats has to generate reports it will look up the parameter value from the specified file and replace it with the associated replacement value. This will happen during report generation, **not** when the log files are processed (so "Use existing database..." can be used as method of database population). The format of the lookup file contents is as follows:

value1=replacementvalue1<LineFeed>
value2=replacementvalue2<LineFeed>
value3=replacementvalue3<LineFeed>
....
....

for instance:

AK=Alaska
AL=Alabama
AR=Arkansas
....
....

End of lines can also be <CarriageReturn><LineFeed> ASCII characters instead of <LineFeed>.

You can configure your CMS system (or create a script) to generate export files on a scheduled basis (best is with the Windows Scheduler) and SurfStats will pick up the new parameter values automatically when reports are generated.

Applicable Reports

Reports that shows page and parameter replacements are:

- *Top Pages and Parameters* report in the *Page Reports* section. With this reports all pages requested by visitors is shown in one report with the page and parameter substitutions.

- Reports in the *Conversion Tracking* reports section. With conversion tracking a set of reports are generated for each page configured as a conversion page, so more detail is displayed.

Page Parameters

What is Page Parameters

Page Parameters (also called query string parameters or two dimensional URL's) are parameters and values specified as part of a URL and affects the way the web server executes the page. For instance in the following URL:

http://www.mydomain.com/productpage.aspx?prodid=15

prodid is a page parameter with a value of 15.

If you use page parameters in .aspx (ASP.NET pages) .asp (Active Server Pages), .cfm (ColdFusion), .php, .cgi, .jsp, aspx, exe, .dll and .cwr dynamic pages, the same page requested with different page parameters will be grouped together by default as requests for one page. This means for example that requests for *training.asp?level=1* and *training.asp?level=2* will be grouped together and reported as requests for the page *training.asp*.

You can set up SurfStats to differentiate between requests for these types of pages with different page parameters, e.g. between the *training.asp?level=1* page and the *training.asp?level=2* page to show page requests as requests to different pages.

To do this, edit the site profile and click on the *General* tab. Check the *Separate parameterized queries to same page* option.

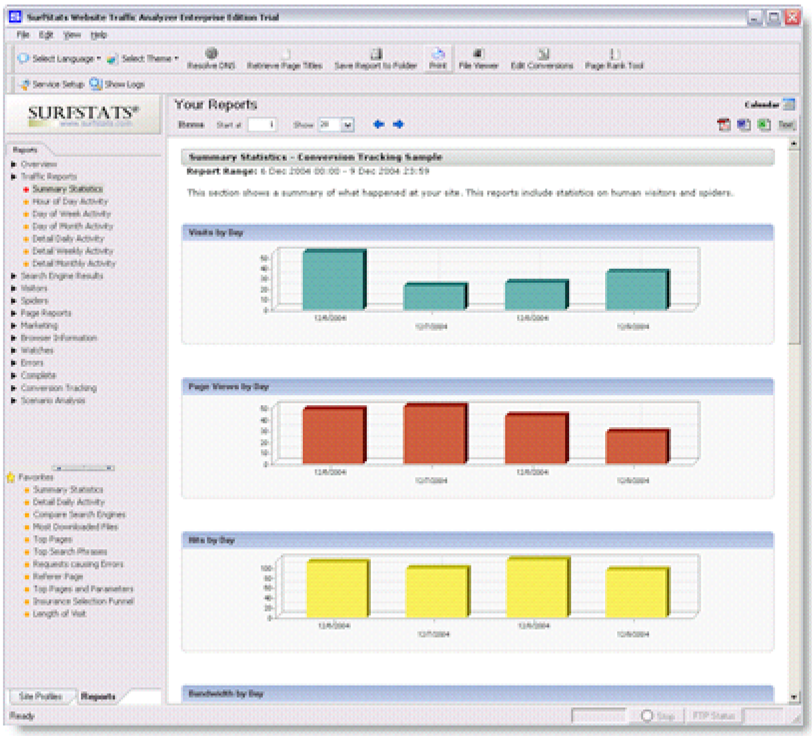
The *Top Pages and Parameters* report shows a breakdown of pages and the associated parameters.

To track how many times a page has been called with a specific parameter you need to define the page as a **conversion page**. Please refer to the section on Conversion Tracking for more information.

Working with Traffic Reports

Internal Report Viewer

The following is a screen shot of the menu structure when reports are viewed from within SurfStats Website Traffic Analyzer:

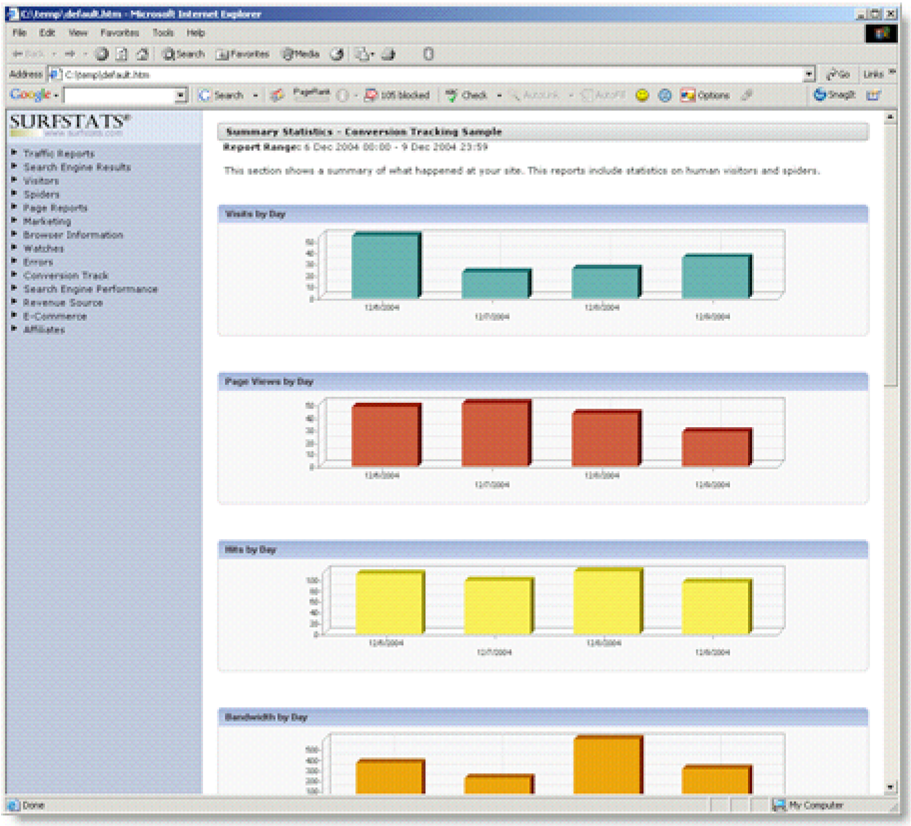


Report category menu items (normally displaying Overview, Traffic, Visitors etc.) are displayed on the left allowing you to choose different types of reports. After the report category has been selected a specific type of report can be selected by clicking a report name on the left-hand side, beneath the report category selected.

New to V8.1 is a favorites section allowing you to specify your favorite reports that you view most often. Drag the selected report name from the top left menu selection to the favorites section at the bottom left to add a favorite report. Report favorites are remembered when the application is shut down and restarted.

Report Viewing by browser

The following is a screenshot of the menu structure when the report destination is set to **Send Report to a Folder**: After SurfStats has generated the reports to a folder you can view the reports by double-clicking the file **default.htm** from Windows Explorer.



All the reports are listed in the left frame and when clicked on, the selected report will be displayed in the frame on the right.

Database Population Method

Select a site profile and click on the Report menu item (or double click the site profile) to generate reports. When one of these actions is performed you are presented with three options to choose from on how to populate the database. The default-selected option would depend on the default selected in Options – Views – Default option for reporting. The three options are:

1. Overwrite database with log file(s) - This option will clear the current data in the reporting database for the site profile, process all the log file(s) and repopulate the reporting database for the site profile database from scratch. Duplicate date/time entries will be added to the database. This means that you can analyze multiple sites and report on them together using this method of database population.

2. Add changes in log files to current database - This option will keep the current database for the site profile and just add new log entries. This option checks for duplicate entries, which will not be added to the database. SurfStats processes the entire list of log files, as it has to check for duplicate entries by comparing entries in the log files with entries present in the site profile's database. Please note that you should not use this option when analyzing log files form multiple web servers or sites as one site. Also note that if you change the profile's filters in Fast mode, the database will be cleared.

3. Use existing database - This option will not process the specified log files and will not add any entries to the existing database. It will use the existing data in the site profile's database.

After database population, the report list and the summary report will be displayed. Reports (except for the summary report) are not generated when reporting to screen until you select the report's menu item on the left. Click the appropriate reporting category tab at the top to open up the list of reports for a specific category on the left.

Reporting Mode

Versions prior to V7.1 featured a Fast and Interactive mode of report generation. Starting at V7.1 reports can only be generated in Fast Mode.

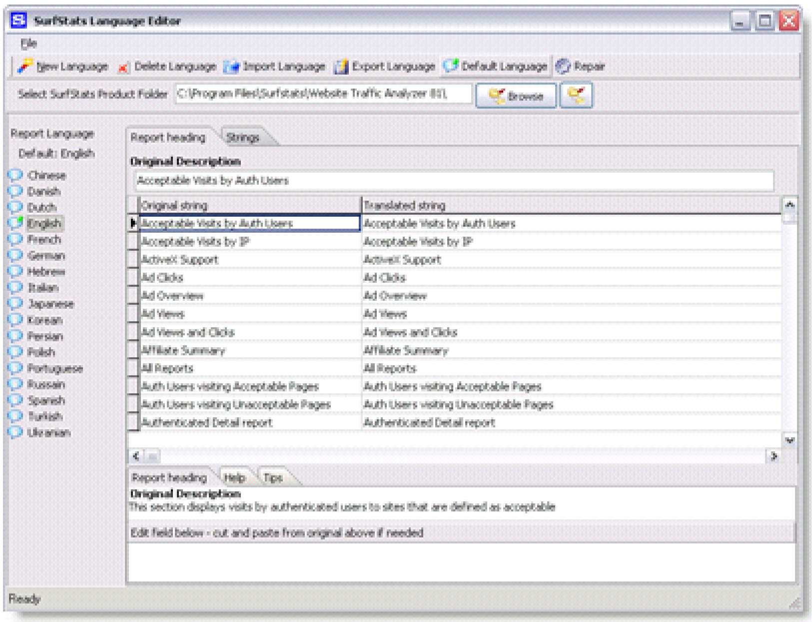
Reporting Language

You can easily add or change the language phrases and terms on reports with the **SurfStats Language Editor** and then select the language to be used on reports from a drop-down list. Western European, Eastern European, Russian and Eastern Asian languages are now supported.

Please note that the current reporting languages in the program were translated by machine and might require editing.

The Language Editor is available when you click on the Tools – Language Editor menu item on the main screen from Site profile mode or by clicking the Language Editor menu item on the left. You can click the Select Language menu item on the toolbar when you are in reporting mode to change the language of the currently generated reports.
You can change the language that a site profile is linked to on the General tab when you edit the site profile.

The Language Editor screen looks as follows:



After the Language Editor menu item is clicked the SurfStats Product Folder has to be selected first. Use the Browse button to select the product folder (default is c:\Program Files\SurfStats<Version>). After the folder has been selected the defined languages will be displayed on the left.

You can change the default language by right-clicking one of the languages on the left and selecting Default Language from the popup menu. If a new language is created it will be based on the default language.

To add a new language, click on the New Language menu item, give the language a name and click OK. Select the newly created language and change the English text to the translated text. You can directly edit text in the right hand panel as well as in the two bottom panels.

To edit an existing language, select the language and change the text. To delete an existing language, right-click on the language and then on the Delete Language menu item on the popup menu.

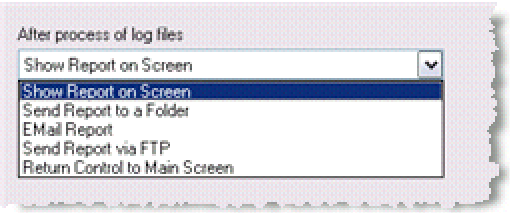
To export a modified or newly created language file, right-click the language and click on the Export Language menu item on the popup menu. Then give it a file name and destination. To import a previously saved/exported language file, right-click a language on the left, then click Import Language and specify the name for the language and file name to import.

The Repair button at the top is for system use only. It is used by personnel from SurfStats to rebuild all language databases after a new term has been added.

Under the tab Report heading, headings to display at the top of reports can be edited. Do not change the Sublink value. Under the tab Strings, strings (or text) displayed on reports can be edited. Do not change the Sublink value.

Report Destination

You can set the destination for reports generated by SurfStats Website Traffic Analyzer.



The report destination can be set using any of the following three methods:

1. Edit the site profile and select the Output To menu item on the left. From the Output to screen you can also select the output format as HTML or Microsoft Word format. Please make sure that grammar and spelling checking is set to OFF in Microsoft Word when generating reports to Microsoft Word. The report that should be the Main Report when SurfStats generates reports for the site profile. This is the report that will get displayed for Summary Statistics.
2. Select a site profile and click on the Report menu item (or double-click the site profile), then on the Results To menu item and select a destination from the After process of log files select box. Changing the selection for this field will also update the selection on the site profile.
3. Set up a schedule and click the Post processing (Click Schedule Task – Add task(s)). Select an option from the After processing log files selection box.

Reports are generated in HTML format and can be read with most browsers, the latest versions of Microsoft® products (like Word, FrontPage, Excel, etc.) and numerous other software.

SurfStats can send reports to the following destinations:

Screen

The report can be viewed in HTML format from within the SurfStats application. Please note that you need Internet Explorer 4 or later for this option. The viewed reports (report to screen) are also placed in the directory:
\Program Files\Surfstats<Version>\Sites\<ProfileName>\Html
where <Version> is the version of SurfStats Website Traffic Analyzer you are running, for instance 8100 and <ProfileName> is the name of the specific site profile.

File Folder

The report(s) are sent to a folder or directory that can be on the Local Machine, a machine on a local or Remote Network (Any Windows Machine) or any Unix or Linux machine running NFS or some form of file sharing system. Note that you can specify sub-directories relative to the system date by using date macros, e.g. the specification:
c:\test\%dd%%Mon%%yyyy%\
would create a sub-directory c:\test\31May2003\ on May 31, 2003 and the reports will be generated in this directory. Note that the directory c:\test\ must exist.
You can create a daily report to a daily folder, e.g. to create yesterday’s date in the format Mondd, use:
c:\%date-1%%mm%/%dd% to create the directory c:\0530 on May 31, 2003
or a weekly report:
c:\Week\%ww% to create the directory c:\Week \22 on May 31, 2003
or a monthly report:
c:\Month\%date-32%%mm% to create the directory c:\Month\April on May31, 2003
Subtract the number of days to safely fall in the previous month if you want to create directories on a scheduled basis, e.g. if you want to run a schedule to create a directory of the previous month on the

fifth day of every month, subtract 6.
Refer to the Date Macro section for more detail.
The site name macro, %site%, is an easy way to get the name of the current profile being processed by SurfStats or the Scheduler. For instance, if you specify the output folder as c:\reports\%site%\ and the site profile name is MySite, the directory c:\reports\MySite\ will be created and the reports will be placed in this directory (note that in this case the folder c:\reports must exist).
Please ensure that you have access rights to the file directory where you wish to write the reports to. If the file directory is a share on a different computer on a Windows system ensure that you have both share level and file level access to the folder. If you use the SurfStats Scheduler to generate reports to a folder (and the Scheduler is started up by the SurfStats Scheduler Service) on a remote share ensure that the user context for the SurfStats Scheduler Service has been set up (Administrative Tools - Services) to a user other than LocalSystem. Also it is recommended to use UNC names instead of mapped drive.

Remote FTP folder

Reports generated by SurfStats can be sent to any machine running a FTP Server (Windows, Windows NT, Linux, UNIX etc.). The FTP Server can be on the local machine, a local or remote network or over the Internet. Also refer to the previous paragraph regarding folder/directory creation.
To set up report generation to a folder on a FTP Server, do the following:

1. Select the site profile, click on Edit Site Profile, click the Site tab and ensure that the retrieval method is set to AutoFTP, then click the FTP tab. Enter the FTP server information and click on Test FTP with FTPClient.
2. Click on Connect to connect to your FTP site and browse to the folder where you want the reports placed. Please note that some servers will only allow you to post to “linked directories” and if not will generate an error “Cannot create folder”.
3. The destination directory will be displayed in the Address box (in this case /mysite- www/test). Highlight the address and copy it to the Windows clipboard (Ctrl – C). Then click on Disconnect and close the application. Paste this address in the Output Directory box.
4. Click on OK and your Auto FTP is set up.

Please note: a common cause of SurfStats not being able to write to a folder on a FTP Server is that write permissions is not set on the folder. Please use a FTP Server administration tool on the server (for instance Internet Services Manager on a Windows FTP Server) to ensure that write permissions is set up.
E-mail
Reports generated by SurfStats can be e-mailed to any valid e-mail address. Reports are first compressed in .ZIP format before being sent.
You need to set up your outgoing e-mail settings before you can e-mail reports. After you have done this, you can then specify the e-mail address to where the reports should be sent.

The e-mail set up panel is located by clicking on the Options menu item, followed by EMail.
You then have to choose your e-mail type, namely SMTP or MAPI. Ask your Network Administrator or refer to your e-mail application for more information.
SMTP - If you have selected this option, you then have to enter the following fields:
SMTP Server - The mail server's host name (required).
Mail From - The sender's e-mail address.
Reply To - The e-mail address where replies should be sent.
Account Name - The sender's account name (if required) on the SMTP Server (if required).
Password - The sender's password (if required) on the SMTP Server (if required).
Note: If the SMTP Server seems to reject sending the mail created by SurfStats please check the relaying options on the SMTP Server.
MAPI - If you have selected this option, you can leave the fields blank to use your system defaults or you can enter the following fields:
Profile - The profile name, e.g. Microsoft Outlook Express.
Password - The password for the selected profile.
Subject - In the Subject field (below the Mail Type field), you can enter the subject that the user who will receive the reports will see. This subject will be displayed on all e-mailed reports.
You can use the %site% macro in the subject field. This macro will display the site profile name in the subject.
Select the site profile, then click on Edit Site Profile, select the Output To tab and select to E- Mail Report. Then enter the recipient's e-mail address. You can also enter these details when creating a new site profile.
You need an Internet connection and be connected to the Internet to e-mail reports.

Please note: As from V7.0 of SurfStats the contents of the Description field of the Site tab will be used as the body text when reports are e-mailed. The reports are zipped and added as an attachment to the e-mail message.
Return control to main screen
You can use this option to populate the database without generating any reports.

Printing reports

The **Print** button on the top toolbar will be available when you are in reporting mode.

To print a single report click on the **Print** menu item to print the report that is displayed in the report viewer panel.

1. To print out all reports, you can:
   Generate reports to a folder and open the default.htm. Then select the Print from your browser and select to print all linked documents or -

2. Right click on the left index page and
   a. Select the **Print** menu item
   b. Select **Only the selected frame** and **Print all linked documents** and
   c. Click on **Print.**

Please note that the application uses the Internet Explorer Control to print reports. This means that if you have problems with printing margins, background printing or if graphs does not print properly, check your print setup in Internet Explorer.

Report format

When generating reports to a folder, FTP or e-mail the Report format can be specified as HTML or MS Word. Version 6.6 and later of SurfStats requires MS Office 2000 or later on the system running SurfStats Website Traffic Analyzer to generate reports in MS Word format.

Report Descriptions

This is a list of reports with descriptions available in SurfStats Website Traffic Analyzer. When you click on a report menu item, the report opens up in the viewer window. Please note that some of the reports listed are not available in FAST mode and some reports require referrer information in log files. Refer to the Glossary of Terms for definitions of terms used in the reports. Also note that the report category tabs and report name menu items are language dependant.

Overview Reports

Overview Reports provide a quick combined, graphical overview of the reports on the other report category tabs. The function of the overview reports is to get a general idea of site activity in a short period of time.

Summary Statistics

This report (on the Traffic reporting category tab) gives a general overview of web activity for the requested time period. Note that averages are calculated by using totals for the specified time period divided by the real value of time, e.g. if the time period was 3 hours, the divider would be 0.125 (3/24). This means that if the total for the three-hour period were 200 clients, the average for the day would be 1600 (200/0.125).

Traffic Reports

The Traffic reports display site activity information for the period for which the reports are generated. This section contains detail as well as summary reports.

The Hour of Day, Day of Week and Day of Month reports provides summarized site activity information for the period being analyzed. This means for instance that if you analyze log files containing three months worth of data SurfStats will add up and group the data for hours 00- 23, Monday to Sunday, day 01-31.

The Detail Daily, Detail Weekly and Detail Monthly reports provides detail site activity information for the period being analyzed. This means for instance that if you analyze log files containing three months worth of data SurfStats will show report data for ninety days for the Detail Daily (or the limit imposed by the maximum allowable number of graph/table entries for the profile), twelve weeks for the Detail Weekly and three months for the Detail Monthly report.

Hour of Day Activity - This report shows the total hourly activity for each hour for the report period. This information is useful to determine for instance what time of day is best to do system maintenance. Note that the report is generated in the system time of the host server and then adjusted by the GMT Time offset (as specified on the General tab for the site profile). Please refer to the section on editing site profiles for more on the GMT Offset.

Day of Week Activity - This report shows the activity for each day of the week for the report period. This information is useful to determine for instance what day of the week to do system maintenance.

Day of Month Activity - This report shows the activity for each day of the month for the report period. This information is useful to determine trends for marketing purposes.

Detail Daily Activity - This report shows the activity for each day of the requested report period. This information is useful to determine trends for marketing purposes.

Detail Weekly Activity - This report shows the activity for each week of the requested report period. This information is useful to determine trends for marketing purposes.

Monthly Activity - This report shows the activity for each month of the requested report period. This information is useful to determine trends for marketing purposes.

Detail Daily Activity - This report shows the activity for each day of each month for the requested report period. This report s a good indication of traffic activity for control purposes and the information is useful to determine trends for marketing purposes.

Views per Visit - This report shows the number of page views requested per client session. This report gives an indication of how compelling your site content is for visitors when they reach your web site. If most visitors view only zero (only downloaded files(s) for instance – did not view a page) or one page you should probably do something to make the contents of your site more appealing.

Tip - This report should be used in conjunction with the Length of Visit report. Keep records of this report to determine if the number of pages increases or decreases due to marketing campaigns. Filter out visits from search engines and their robots to determine "real visitors".

Length Of Visit - This report shows the duration (in minutes) of visits to your site. This information is an indication of how compelling your site is to visitors. The average US user spends less than one minute on a site. Strive to increase the time spent on site.

Tip - This report should be used in conjunction with the Views per Visit report. Keep records of this report to determine if the time spent increases or decreases due to marketing campaigns. Filter out visits from search engines and their robots to determine "real visitors".

Visitors and Spiders

Visitor reports display site activity information about visitors to your site.

Who is on – This report displays which users accessed the web site in the past 6 hours. If you are analyzing log files older than 6 hours no data will be displayed in this report.

Recent Spiders – This report displays which spiders (from Search Engines) accessed the web site in the past 6 hours. If you are analyzing log files older than 6 hours no data will be displayed in this report. A visitor is marker as a spider (or search engine robot) when it tries to access the file called robots.txt on your web server.

Top Visitors - This report shows the IP address and/or the Whois/domain name and their relative

activity level on the site. SurfStats Website Traffic Analyzer does not differentiate between different users using the same IP Address. However it can group together visits from different IP addresses for a user with the same cookie (see section on Global Settings).
Note that this report excludes search engine robots.

Top Spiders - This reports shows the most active search engine robots that visited your site.

Top Visitor Organizations - This reports shows the most active organizations that visited your site. This is determined from the domain name of the visitor. Spiders are excluded from this report.

Tip - You must resolve IP Addresses or have the Visitor Origin Database installed before requesting this report.

Top Spider Organizations - This reports shows the most active spiders (search engine robots) that visited your site. This is determined from the domain name of the visitor.

Tip - You must resolve IP Addresses or have the Visitor Origin Database installed before requesting this report.

Top Visitors by Bandwidth - This report shows the top visitors sorted by bandwidth. Spiders are excluded from this report.

Top Spiders by Bandwidth - This report shows the top spider visits sorted by bandwidth.

Authenticated Top Visitors – This report shows detail activity of authenticated users accessing files and pages that require authentication (a login user name and password).

Top Visitor Countries - This reports shows the countries where your client’s domain names are registered. Please note that it may not be the actual location of your client. Some Website Traffic Analyzers assumes that .com and .net domains are located in the US but this is not always the case. Spiders are excluded from this report.

Top Spider Countries - This reports shows the countries where spiders that visited your site came from. Please note that it may not be the actual location of the search engine. Some Website Traffic Analyzers assumes that .com and .net domains are located in the US but this is not always the case.

Top Visitor Regions - This reports shows the regions where your client’s domain names are registered. Please note that it may not be the actual location of your client. Spiders are excluded from this report.

Top Spider Regions - This reports shows the regions where the spiders that visited your site are located. Please note that it may not be the actual location of the search engine.

Page Reports

The reports on the Pages tab display details about pages and files that were accessed or downloaded by visitors to your web site. Depending on the extension of a file it will be classified as a page (for example .htm, .asp, .cfm etc), file (for example .exe, .zip etc), image file (for example .jpg, .gif etc.) etc. SurfStats Website Traffic Analyzer uses what is specified in Options – File Types to classify file types as pages (documents), files, audio, video or images.

Most Downloaded Files - This reports shows the most downloaded files. The graphs show the number of download request. Note that this is not always equal to the number of downloaded files, as users might have used download managers like GetRight and GoZilla for instance. A resumed download would be counted as a new download request.

Tip - You should also look at the Bandwidth figure and divide it by the file size to get a reasonably accurate number of downloaded files.

Top Pages - This reports shows the page names that was requested successfully, the most times by visitors to your web site. This report shows pages that were requested that didn’t contain a ? character in the name (indicating a static page).

Tip - Use filters to display hits to specific pages and then also look at the Length of Visit report.

Top Pages and Parameters – This report shows pages that were requested that did contain a ? character in the name (indicating a dynamic page with query parameters).

Top Entry Pages / Files - This report shows which pages or files were used to enter the site. This report should be used in conjunction with the Top Referrer Sites report. The information is useful to determine how visitors came to your site.

Top Exit Pages / Files - This report shows the last file or page that a user accessed before leaving your site. It is loosely called an exit page or file.

Top Images - This report shows the most requested images. Note that you can add or delete image types to track in the filters as well as in the option specification. You can restrict the number of requested images to be displayed in tables and graphs. Note that if you have selected to filter out hits to image files (which is selected by default for a profile) no images will be displayed here.

Top Audio Files – This report shows the most requested audio files.

Top Video Files – This report shows the most requested video files.

Top Unknown Files – This report shows the most requested unknown files. A file is classified as unknown if the extension type cannot be classified as a page, file, audio, video or image file (as set up in Options – File Types).

Site Bookmarked - This report shows which of your directories has been bookmarked by users of Internet Explorer 5 and Netscape 6 or higher. It is an indication of how interesting visitors find your site. It shows the directory and the number of hits to the favicon.ico file.
Tip - This report should be used in conjunction with the Error Referrer report. Keep records of this report to determine if the number of times your site were bookmarked increases or decreases due to marketing campaigns. To enable this feature in you logs, place a "favicon.ico" file in every directory visitors have access to. IE5 will then use this instead of its default icon. For more information, visit http://www.favicon.com Also filter on "*favicon*" to see other reports with information on hits to this file.

Single Access Pages / Files - This report shows which pages were accessed only once by visitors to your web site. It is an indication of how compelling visitors find your web site. Search engine robots usually generate many of these single page requests.
Tip - This activity should be tracked over time to determine if this number increases (the contents of your site becomes more compelling) or decreases due to marketing campaigns. Use filters and also compare to the Site Activity reports.

Most Downloaded File Types - This report shows the most downloaded file types (grouped by file extension). Note that you can add or delete file types to track in the options section.

Most Accessed Directories - This report shows the most accessed directories. It gives an indication of which areas of a site gets accessed the most often. Hosts can use this report to determine their clients' relative web site activity.

Browser Information

Browser reports show detail of the types and versions of web browsers that visitors used to access your web site.

Top Browsers - This report shows the most used browsers that accessed the site. This information is useful to determine if the majority of visitors will see the site as intended, e.g. Internet Explorer users will be able to see SurfStats reports with shading and Netscape users will not. Note that you can add new browser footprints by selecting Options - Browsers (in Site profile mode).

Netscape Browsers - This report shows the most used Netscape browsers that accessed the site.

Internet Explorer Browsers - This report shows the most used Internet Explorer browsers that accessed the site.
Wap Devices - This report shows the most used WAP devices that accessed the site.

Visiting Spiders - This report shows the spiders that accessed the site. Note that you can add new spider footprints by selecting Options - Browsers (in Site profile mode).

Download Managers - This report shows the download managers that were used to download files from the site. Note that you can add new download manager footprints by selecting Options – Browsers (in Site profile mode). This information can be used to provide a better service.
Tip - Ensure your site supports resumed downloads if you host large files for users to download.

Top Platforms - This report shows the platforms or operating systems used by visitors to the site. This information is useful to find out more about your target market. The current list of platforms detected is:

- Windows 3.X/95/98/ME/NT/2000/2003/XP/2008/20127/8
- Win32
- Windows CE
- Linux
- Unix
- HP Unix
- Xwindows
- OS2
- Mac
- Mac PowerPC
- SunOS
- WebTV
- PalmOS

The platform is listed as unknown when none of these Operating Systems are found.

Capabilities
Browser capability reports lists capabilities of web browsers that visits your site.

Frame Support - This report shows if the visitor's browser has frame support. It is useful to decide whether to use Frames on a site.

VBScript Support - This report shows if the visitor's browser has VBScript support. It is useful to decide whether to use VBScript on a site.

JavaScript Support - This report shows if the visitor's browser has Javascript support. It is useful to decide whether to use Javascript on a site.

Java Applet Support - This report shows if the visitor's browser has Java Applet support. It is useful to decide whether to use Java Applets on a site.

ActiveX Support - This report shows if the visitor's browser has ActiveX support. It is useful to decide whether to use ActiveX controls on a site.

Search Engine Results

Search Engine Results reports shows which third parties referred visitors to your site. If these reports are empty it is likely that your web server has not been set up to log referrer information. Please refer to the section on Working with Log Files for more on this.

Top Referring Sites - This report shows the domain that referred a visitor to a page or file on your site. This information can be used to see which sites link to your site and to track banner ads that you have placed on other sites. Note that Unknown referrer means that visitors either typed in the link to the site or used book-marked links.

Tip - You can use your URL in the filters to show only internally referred hits or those from external sites.

Top Referring Domains - This report shows the URL that referred a visitor to a page or file on your site. The referrer link can usually be followed back to the referrer's page. This information can be used to track banner ads that you have placed on other sites.

Tip - You can use your URL in the filters to show only hits referred internally or those from external sites.

Top Search Engines - This report shows the Top Search Engines that referred visitors to the site. Note that you can add or delete search engine footprints (Options – Search Engines in Site profile mode).

Top Search Phrases - This report shows the top search phrases visitors used to find the site on search engines. This is useful to see what phrases visitors use to find your site. Use this report to know what keywords to add to the META tags in your web pages.

Top Search Keywords - This report shows the top keywords visitors used to find the site on search engines. This is useful to see what words visitors use to find your site. Use this report to know what keywords to add to the META tags in your web pages.

Search Phrases Per Engine - This report shows the search phrases used per search engine and is useful to see your relative visibility for search phrases in search engines. You can then use this information to create new pages with search phrases that you want to be found on by a specific search engine. Note that you can set the number of records to be displayed in tables.

Marketing Reports

These reports display results from ad tracking.

Ad Views - This report shows the number of times that a banner (graphic) ad was viewed. This report is useful when selling advertising space on a site.

Tip - Create a special directory (e.g. \ads) on your web site in which to place your ad images. This makes it easier to update and filter on.

Ad Clicks - This report shows the number of times that a banner ad was clicked on. This report is useful when selling advertising space on a site. Please note that banner ads usually link to external files and clicks on external links normally do not generate hits on your server's log file. To create a hit to external files would require a hit to a file on the server, which then redirects the user to the external link. Also see the Advertisement section on how to create hits in log files from clicking on links to external files.
Tip - Create a special directory (e.g. \redir) in which to place the redirect files. This makes it easier to update and filter on.

Errors - Error reports shows errors that were generated by your web server.

Server Errors - This report shows the range 5 or server errors, e.g. 501, which means Not implemented.

Summary of HTTP Errors - This reports shows a summary of the HTTP errors on the site, e.g. 404 - File not found.

Requests causing Errors - This report shows the error codes, error description, file that caused the error and the number of errors from visitors' browsers accessing the server. This information is useful to identify and rectify errors on your site.

Detail Daily Errors - This report shows the errors codes sent to visitors' browsers accessing the server on a daily basis. This information is useful to identify errors trends on your site.

Error Referrers - This report shows the referrer link that caused errors from visitors. This helpful to identify and rectify invalid links to and on your site.

Complete

You can generate all proxy reports to folder by selecting File – Save Report to folder. Then use Windows Explorer to browse to the folder and double-click the file default.htm.

Reports not in SurfStats

Path though site – If an organization designs the layout of a store it is most often designed in such a way that the most needed items are usually placed the furthest from the store's entrance. This is done so that clients have to walk through aisles containing special offers that entice them to buy other products. This concept is probably what enticed some Website Traffic Analyzer software developers to come up with the "path through site report". An Internet store differs from a normal store in the sense that the Internet visitor can follow multiple paths at the same time.

The creation of a "path through site report" is extremely memory and space intensive and slows down the parsing and report generation process significantly. This makes it infeasible to implement in a log analyzer with a database.

The only useful information you get from a "path through site" report is the basically the same that you can get from the Entry Pages and Most Requested Pages as well as the Top Destinations Reports, except that the files are not shown in sequence. It is also not of much use to determine the correct path if a visitor opens more than one browser window of your website randomly.

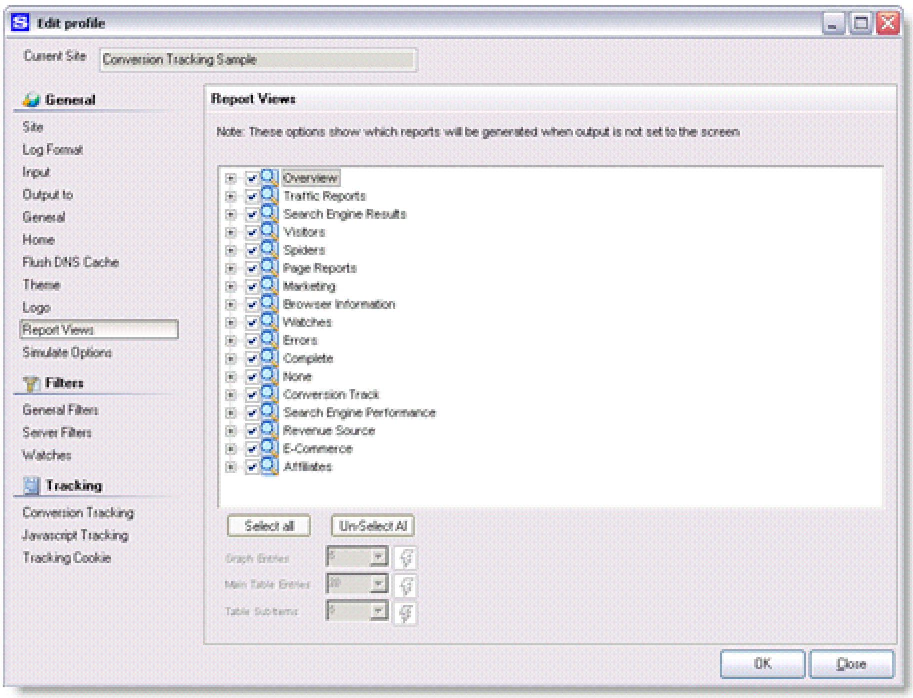
To see all the pages/files that a specific visitor accessed, you can use the Include Domain/IP like filter.

The Conversion Tracking Reports in general provide a lot more useful information than a Path through site report.

Server attack – This report, which used to be in versions 7.0 and earlier had been removed.
US States - This report is based on where the domain name of the visitor is registered or deducted from the suffix of the domain. This report is not accurate and of little use due to many inaccurate assumptions and very few US States suffixes in visitors' resolved names.

Specifying Reports

You can specify which reports should be generated for a specific site profile. Select **Edit Site – Report Views**.



You use the panel to select the report types to be generated and the number of records to be displayed in the graphs, main tables and sub-items in tables. Please note that the number of graph entries are limited to a maximum of 20. Click the yellow lightning button to apply the change.

The Professional and Enterprise Editions, V7.1 and later, features a **View set Editor**, which makes it easy to configure view information for multiple site profiles.

Resolving IP Addresses

To find a server on the Internet, you would have to specify it's unique URL (Uniform Resource Locator). The Uniform Resource Locator is a means of identifying an exact location on the Internet. An URL consists of four parts:

Protocol Type (http, ftp etc.),
Domain name,
Directory Path (/), and
File Name (default.htm),

The domain name can also be replaced by an IP Address e.g. 64.176.57.184, which is the IP Address for surfstats.com. So if you want to reach the web page default.asp in the surfstats directory on SurfStats's web site you can (in the Address field of a browser) either type:
http://www.surfstats.com/surfstats/default.htm or you can type
http://64.176.57.184/surfstats/default.htm

An IP Address consists of four groups of numbers, separated by dots (for example 64.176.57.184). A server on the Internet must have an IP Address but does not need to have a domain name. The IP Addresses and corresponding Domain Names, if it exists, are stored on computers on the Internet running a service called Domain Names System (also called DNS servers).

The Internet Corporation for Assigned Names and Numbers or ICANN, governs and coordinates the distribution of IP addresses. ICANN delegates IP address allocation duties currently to three Regional Internet Registries (RIR):

ARIN for the Americas and Sub-Saharan Africa at http://www.arin.net/
RIPE NCC for Europe, Northern Africa, and the Middle East at http://www.ripe.net/
APNIC for the Asia-Pacific region at http://www.apnic.net/

These RIR sub-allocate ranges of IP addresses to local Internet registries and large ISPs, which then sub-allocate blocks to smaller service providers or assign addresses directly to end users. This hierarchical system ensures global uniqueness and distributed management for IP addresses, just as the DNS does for domain names.

A DNS lookup or IP Address Resolution refers to the querying of a DNS server for the corresponding domain name or IP Address. Forward DNS lookup is using the Internet Domain Name to find an IP Address. Reverse DNS lookup is using an Internet IP Address to find the corresponding Domain Name. The DNS query might be to the local DNS server at your service provider, some intermediate name server, or the root DNS server.

The IP Address of a visitor to your site is logged in the log file and it is therefore possible to get the domain name of the visitor from this address. DNS

resolution can take a long time and is dependant on the number of IP addresses to resolve, the congestion of the Internet, the availability of a Reverse DNS record (called a PTR record) for the IP Address and DNS servers.

You can find out more about the process of DNS lookup on the Internet at:
http://whatis.com/forwarddns.htm
http://www.amnesi.com/
http://www.whoix.com

The best way to get the resolved IP Address names is to set your server (or ask your host) to set up reserve IP lookup for your log files. The resolved names will then be placed in the log files and can then be used with Surfstats Website Traffic Analyzer.

Geographical mapping

You can determine the geographical location if you know your visitor's Domain Name or URL if the suffix in the URL contains the type of organization and the country. Also refer to the glossary and ISO

Internet and Country Codes.

Some US domain names has the city and state as a suffix in their URLs. A client report can be generated showing the number of clients from certain US cities and states. Unfortunately the number of domain names that has this suffix is so small that it does not event warrant the effort to generate such a report.

Reports will show that AOL users are coming from somewhere in Virginia because of AOL's use of proxy servers to handle their web requests. Thus, it is impossible to attach meaningful geographic location data to an AOL IP or other large ISP's that use proxy servers.

Most ISPs and networks use some form of dynamic IP allocation for its users. This means that a user might have a different IP address every time he or she logs on to the Internet. You can see how this might affect the accuracy of a geographical report.

Visitor Origin Database

A separate IP Address database, based on Whois queries is now available as a separate product from SurfStats Website Traffic Analyzer. Pre looked-up entries from this table are used in reports of the Visitor section if the Visitor Origin Database is installed. Using this database makes DNS lookups obsolete. It is much faster and more accurate than Reverse DNS lookups. This Visitor IP Address table is about 30 MB compressed or about 150 MB uncompressed.

The Visitor Origin Database is downloaded and installed separately and registered separately. When you purchase the Visitor Origin Database you will receive the necessary details from SurfStats.

Please note: The entries in the Visitor Origin Database cannot be edited or viewed. It will only display on reports.

DNS Setup

Click **Options – DNS Settings – DNS Lookup**.

We recommend that you do not adjust the number of threads to use to more than 100 if you have a slow connection to the Internet or if you are going to do lookups from behind a firewall.

DNS Ranges

A feature introduced in V6.6 of SurfStats Website Traffic Analyzer is the DNS Ranges database that contains ranges of IP Addresses. This feature can accelerate the DNS lookup process. The DNS Ranges database contains a range of IP Addresses. You can add to this IP Range database (**Options – DNS Settings – DNS Ranges** in Site profile mode – Check **Allow edit**, click on the list and press Insert). On the DNS Ranges list you can also add the address range of your local Intranet (which no external DNS server will be able to resolve anyway). For instance you can add the following entry if your organization (called mycompany located in the US) has an Intranet IP Address range from 192.168.1.1 to 192.168.1.255:

Clicking the **Lookup** button will connect to arin.net with your web browser and try to resolve the Start IP address of the selected row.

SurfStats Website Traffic Analyzer versions prior to V7.0 came with a lot of pre-resolved entries in the list of DNS Ranges. For V7.0 and later most of these entries are located in the optional Visitor Origin Database. After installing V7.0 the list of DNS Ranges should be empty.

Clearing the DNS database

Resolved IP Addresses are stored in the DNS Cache/Database . This file can become large and can be cleared or flushed.

DNS Lookup through proxy server

To enable DNS lookups through a proxy server, set up the proxy server or firewall to allow the following TCP/IP and UDP/IP traffic **out** from the Intranet to the Internet:

- UDP: Port 53 and 137
- TCP: Port 53

Resolving Sequence

SurfStats will resolve IP Addresses in the following sequence when lookup values needs to be displayed in reports:

- The IP Addresses (DNS Cache) database will be queried.
- The DNS Range database will be queried.
- The Visitor Origin Database (if installed) will be queried.
- The DNS Server (as set up in your TCP/IP properties) will be queried with a reverse DNS lookup if reverse DNS lookups are requested.

The first of the above to provide a lookup value for the IP address will be used in reports.

Scheduling tasks

The SurfStats scheduler is used to schedule the automatic generation of website traffic reports. You access the SurfStats scheduler by clicking the **Schedule Tasks** menu item (in Site profile mode). When the Scheduler is activated the Scheduler window is displayed.
A scheduled task is associated with a specific site profile. You can set up multiple schedules for a specific site profile.
Note that the Active checkbox for the schedule must be enabled for a scheduled task to run.
You can toggle the Active status by clicking the Activate/Deactivate menu item when a specific schedule is selected (the Status column does not display anything).
Please note that errors in log files will be ignored when report generation are scheduled.

Overview

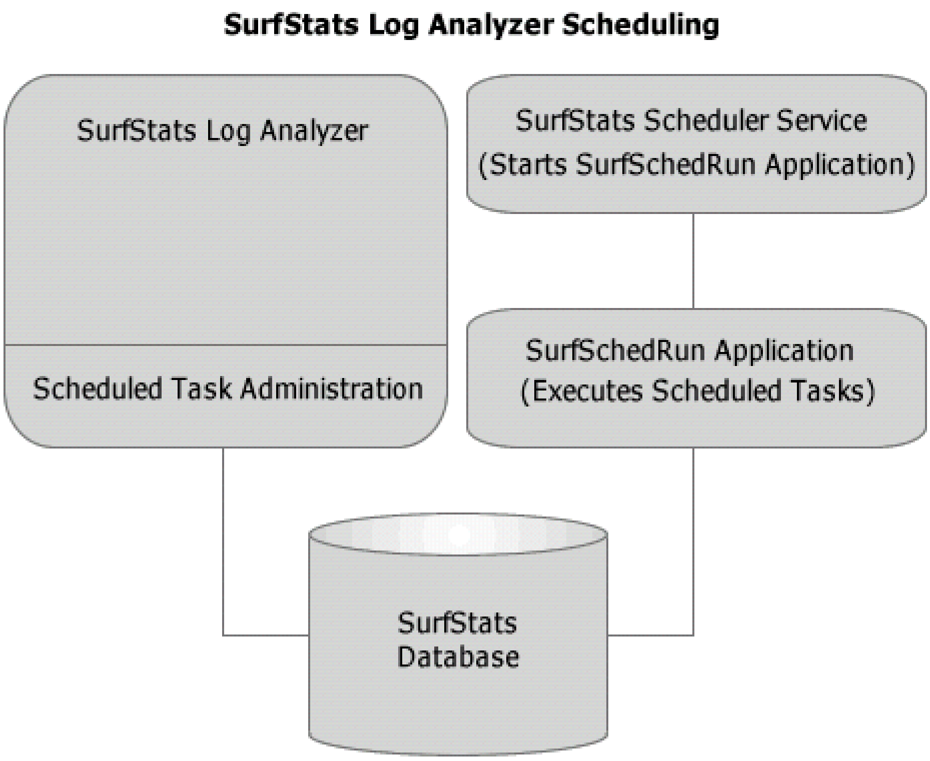
Scheduling in SurfStats Website Traffic Analyzer V7.0 and later consists of the following three components:

1. **Scheduling Task Administration** done from within the SurfStats Website Traffic Analyzer application (clicking Schedule Tasks).

2. The **SurfStats Scheduler Service**, which has the sole responsibility to start the **SurfSchedRun** application if the service itself is running and the **SurfSchedRun** application is not running. The **SurfStats Scheduler Service** is a Windows service registered in the Windows Services Database (Control Panel – Administrative Tools - Services).

3. The **SurfSchedRun** application that scans the list of scheduled tasks and executes them. When the **SurfStats Scheduler Service** starts the **SurfSchedRun** application the activities of the **SurfSchedRun** application can be seen in the status bar when selecting **Schedule Tasks** menu item from within the SurfStats Website Traffic Analyzer application. The **SurfSchedRun** application scans the list of tasks, decide which to execute, executes them and then shuts itself down, to be started up later again by the **SurfStats Scheduler Service**. This is done to boost scheduling reliability.

The diagram below illustrates the relationship between the SurfStats Website Traffic Analyzer Scheduling components:



Please note: You need a Windows Operating System that supports Services to make use of SurfStats Scheduling.

Administering scheduled tasks

Setting up scheduled tasks is performed from the SurfStats Scheduler. The Scheduler Application can be accessed by clicking the **Schedule Tasks** menu item from within SurfStats Website Traffic Analyzer.

Click the Add a task menu item to add a scheduled task or select a scheduled task and click on the Edit a task menu item (or double-click the scheduled task) to edit a scheduled task.

Select the site profile (when adding a task) that this specific task is associated with from the list box on the left-hand side. You can *select multiple site profiles* by using Ctrl or Shift with the mouse and create multiple schedule tasks quickly. The tasks will be completed one after the other.

On the **Run schedule at** tab you can select the start date, start time and repetition interval for the schedule that you wish to edit or create. Clicking the Today or Now button will fill in the current date or time in the Enter start date or Enter start time field. The Repeat every field can be set to specify how often the task should be executed.

On the **Report format** tab you can select the report format (HTML or MS Word). Reports will be generated in the format that you specify here. Note that in order to generate reports in MS Word format you need MS Word installed on the system that you are running SurfStats on.

The version of the report files generated will be in the same version as the version of MS Word that you have installed.

On the **Period** tab you specify the period for which reports should be generated. You can specify a period or select all the data from the log files (Complete Log). If you choose to specify a period the Scheduler will create the start and end date and time for the report automatically when the schedule is run depending on the options selected.

On the **Post Processing** tab you set processing options to:

1. Rebuild database completely, i.e. clear the database and add all the log file entries to the database.

2. Add changes in log files to database, i.e. add only changes in the log file to the database or
3. Do not process log files, i.e. do not add the log file information to the database and:.

After processing of log files, you can select to:
1. Send report to a folder. Tip: If you are generating reports to a folder on a remote computer use a UNC name (for instance \\remotepc\reports\MySite), also specify a valid account with sufficient right (not LocalSystem) for the SurfStats Scheduler Service (Administrative Tools - Services).
2. Email report. Please note that this setting overrides the e-mail destination set when you created the site profile.
3. Send report via FTP
4. Return to main screen.

You can use the %site% (for the profile name) and the date macro wildcards to generate output folder names at runtime.

Note that if you select Send report to a folder you should specify different Output Folders for different schedules to prevent reports from being overwritten by reports generated by a subsequent Schedule.

Click the Go button after when you have filled in all the necessary detail for a task. The task(s) will be added to the list of tasks.

You can set up scheduled tasks to start at the same time. The tasks will then be completed in sequence.

The Service menu item is used to install, start, uninstall and stop the SurfStats Scheduler Service in the Windows Services Database.
The Activate/Deactivate will toggle the Active status of the selected task. The SurfSchedRun application will only run Active tasks.

The View Log menu item will display scheduling activity. This is useful for debugging. To print the log from the View Log screen click the Print log button. Note that you need to have one printer selected as the default printer to print the log. If you have not selected a default printer close the Scheduler, select a default printer from Printers in Control Panel and restart the Scheduler.

Creating multiple schedules

You can create multiple schedules for one or more site profile from the **Add task** screen. To do this:

Specify the schedule information

Select the site profiles that you want schedules for (with Ctrl-click r Shift-click)

Click on the GO button.

Windows service

Running as a service means that SurfStats Scheduling can be active without the need for a user to be logged into the computer on which SurfStats Website Traffic Analyzer runs.
To start or stop the service, click the **Service Options** menu item on the main toolbar of the SurfStats Scheduler

The functions of the buttons are:

Install service - Installs the service and also makes the **SurfStats Scheduler Service** visible and accessible from the Windows Control Panel.

Uninstall service - Uninstalls the service from the Windows NT/2000/2003/XP service database. Note that if the service was installed and started, you need to stop and uninstall it if you do not want to start it each time after a reboot (or you can change the Startup type of the service in the Windows Control Panel).

Stop service - Stops the SurfStats Scheduler service. Note that you have to uninstall the service if you do not want it to start on a reboot (or change the Startup type of the service in the Windows Control Panel). The function of the SurfStats Scheduler service is to start the **SurfSchedRun** application (if it is not already running).

Start service - Starts the service. When you exit the program and reboot the computer when the service is running, it will be started automatically after the reboot without the need for a user to log in. Note that the service will run in the context of the Local System account. You can change the user context from the Windows Control Panel if required. This should be done if SurfStats Scheduling needs to retrieve log files or generate reports to a folder on a remote computer.

The information box beneath the buttons displays the status of the service, i.e. The service is running/not running, etc.

For one or more scheduled tasks to run the following is required:
- A Windows Operating System that supports services (for instance Windows NT4, 2000/2003/XP Professional).
- Task(s) must be scheduled.
- The task(s) must be **Active**.
- The **SurfStats Scheduler Service** must be installed and must be running.
- Report output must not be set to screen.

The scheduler checks for scheduled tasks to run every minute.

You can use **View log** to view the schedule log to see if scheduled tasks have been completed

Running a Scheduled Task

Versions of SurfStats Website Traffic Analyzer prior to V7.0 featured a **Start Now** menu item in the Scheduler application. From V7.0 the only way to run a task is to set the date/time of the task, ensure the task is active and ensure the **SurfStats Scheduler Service** is running.

The **Start Now** feature was removed from V7.0 because the administration of scheduled tasks is now separate from the running of scheduled tasks. The scheduled task will be started by the SurfStats Scheduler Service at the correct time.

Saving and Restoring Data

Site profiles, schedules and other global data are stored in the SurfStats database. Make a backup copy of your setup by clicking on *File - Backup*.

Database

Do not use the SurfStats database or the installation folder sub folders as an archive for your log files. Compress your log files and store them in a different location and use that as an archive .

The database is a temporary storage to simplify report generation. The database tables cannot be saved and restored. The database consists of complex interlinked tables and if changes are made to the structure, it is a nightmare to migrate data from a previous version.

The program parses a log or hit line into different fields, then applies filters (where applicable) and then enters the records into the applicable tables in the site profile's database. The site profile database is used as the source of information from which reports   are generated.

if you use the second option of database population (add new entries in log files to the database) and make changes to filter settings, the resulting reports would not be accurate. This is because existing entries present in the database are not affected by changes in filter settings.

Whenever you make changes to filters it is also recommended that the first method of database population (overwrite database) is used to ensure that filters are applied correctly.

Saving the DNS database

The IP Addresses and the resolved domains (also for intranets) are stored in the *IPMain.dbf* file. You can view and edit this file by clicking on Options - DNS Settings – DNS Ranges.

You can also export and import it in this section - click on Export and save the file. To restore the file, click on Import and select the file to be restored. Here is a description of the file layout:

Row Def.
1 Start IP as Number*
2 End IP As Number*
3 Start IP (no wildcards)
4 End IP (no wildcards)
5 Domain
6 Not Used
7 City - can leave blank
8 State - can leave blank
9 Root (net, com, us, etc.)
----
* Calculate as follows:

If IP address is a.b.c.d then Value = (256^3*a) + (256^2*b) + (256*c) + d
Keep a backup of this list by exporting and saving it.

Saving and Restoring profiles

Pre Version 7.0 - Note that with V6.0.1.1, V6.6 and V6.7 you can keep the site profiles intact when installing a new version. Save your current profiles by clicking **File – Save Profiles**, and then specify a filename. V6.6 and V6.7 can only save your current profiles, not schedules or global options. For V6.0.1.1 – V6.5 beta make a backup copy of the surfsite file in the Windows Installation directory and use this as the site profile backup file.

Versions 7.0 and later – Version 7.0 and later can save/restore site profiles, schedules and global options to/from a XML file created by using the **File – Backup** feature from a V7.0 or later installation of SurfStats Log Analyzer. All profile, schedule and global information are stored in a single file that can later be used in a restore (with **File – Restore – V7.0 onwards**). This feature is only avalailable in the Professional and Enterprise Editions.

Site Profile Folder Structure

The following folders are located under the SurfStats installation folder (normally \Program Files\SurfStats for V7.0 and prior, \Program Files\SurfStats <Version>\ for V7.1 and later):

\Data – This directory contains the empty tables for a site. When a site profile is created and a report is requested for the first time, these files are copied to the \Data directory for the created site profile.

\Datamain – This directory contains the shared tables for the site profiles like the IP Address, Country Codes, File Types, etc.
For every site profile created, a new directory (e.g. MySite) is created under the directory \Program Files\SurfStats\Sites\. For the MySite example, created on Drive C, the directory created would be:
C:\Program Files\SurfStats<Version>\Sites\MySite\
The following directories are located beneath the \MySite directory

\Data – This directory contains the database tables for the site.

\Extract – Temporary directory – used when a zip or other compressed file is processed. The files contained in a zip file are extracted to this directory.

\HTML – This directory contains the html reports created

\LogFiles – This directory contains the files retrieved from remote servers using HTTP or FTP (only files from last retrieval).

\Unprocessed – This is the working directory where log files are uncompressed to prior to analyzing. Files download by AutoFTP are placed in this directory prior to parsing and database population. The log files are moved to the \LogFiles directory after processing.

Note: Please do not copy your log files to folders in the SurfStats installation folder, rather store them in a folder created by yourself, for instance c:\logs.

Data Mining

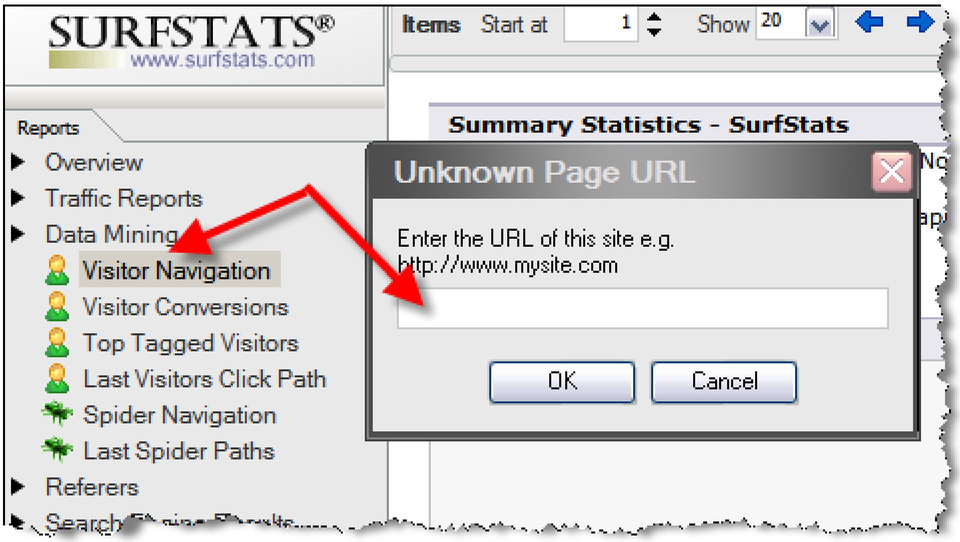
Data Mining is defined as *the automated extraction of hidden predictive information from databases*. SurfStats' powerful Data Mining features makes it easy to recognize and track visitor patterns in your website traffic.

Data Mining is the process of searching large datasets or databases for useful patterns, which normally will not be noticed. A set of reports is available in the Data Mining reporting category for data mining purposes.

Double click on the site profile to parse the log files and open reports.

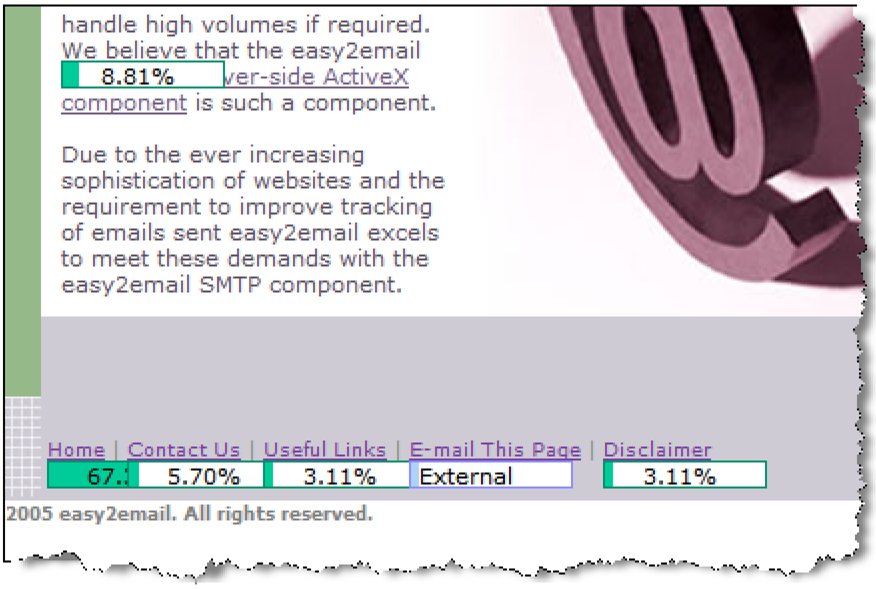
Select Data Mining – Visitor Navigation

If it is the first time you are running the reports for the site profile, you will be prompted to enter the site's URL



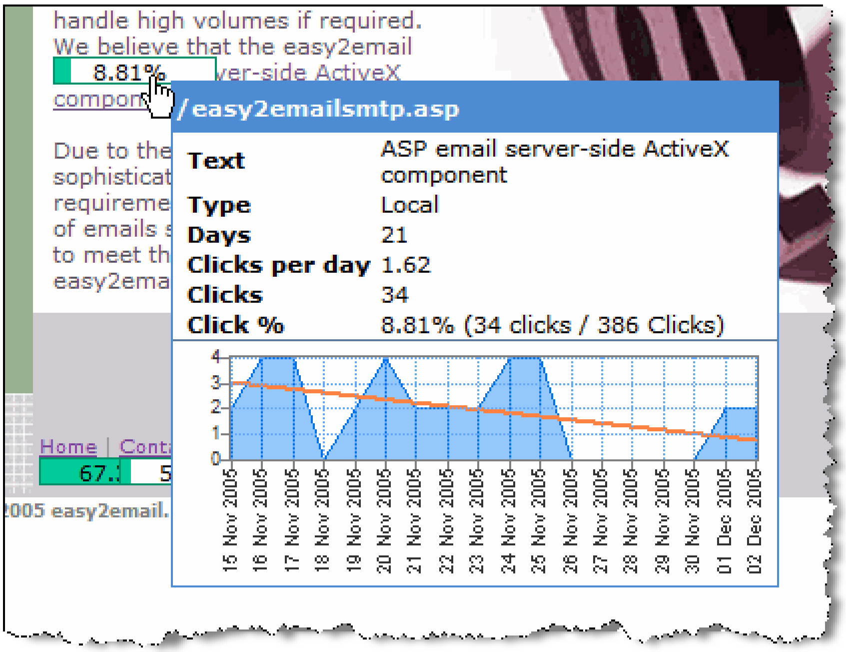
Make sure that you are connected to the Internet and fill in the Site Profile's domain.

The default page of the site will be displayed in the report window and the hit details will be superimposed on the page. Hits on the page to local files will be shown in Green and hits to external files will be shown in Blue.



Note! You have to install the external hits tracker javascript code in every page that you want to track. This javascript creates a hit line in the server log files.

When you move the mouse over the link, a detail graph will pop up to show more detail.



Visitor Tagging and Scenario Analysis (Conversion Tracking)

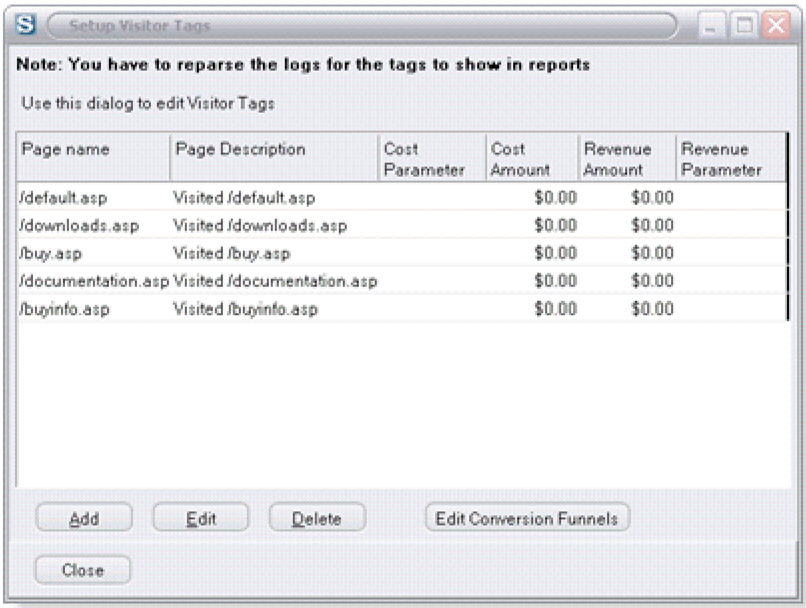
What is Visitor Tagging?

Visitor tagging is the replacement for conversion tracking that was available in previous versions. Visitor tagging improves on Conversion Tracking by allowing site visitors to be grouped into different groups by tagging them, for instance which visitors reached a certain goal page, which visitors entered the site at a specific page etc.

Setting up Visitor Tagging

Visitor Tagging is available in SurfStats Website Traffic Analyzer V8.2 and later. Click the **Edit** link in the **Goals** column for the profile or click the **Visitor Tagging** button on the top toolbar while a specific profile is selected.

This will bring up a dialog screen where you can define conversion tracking settings for the selected profile:



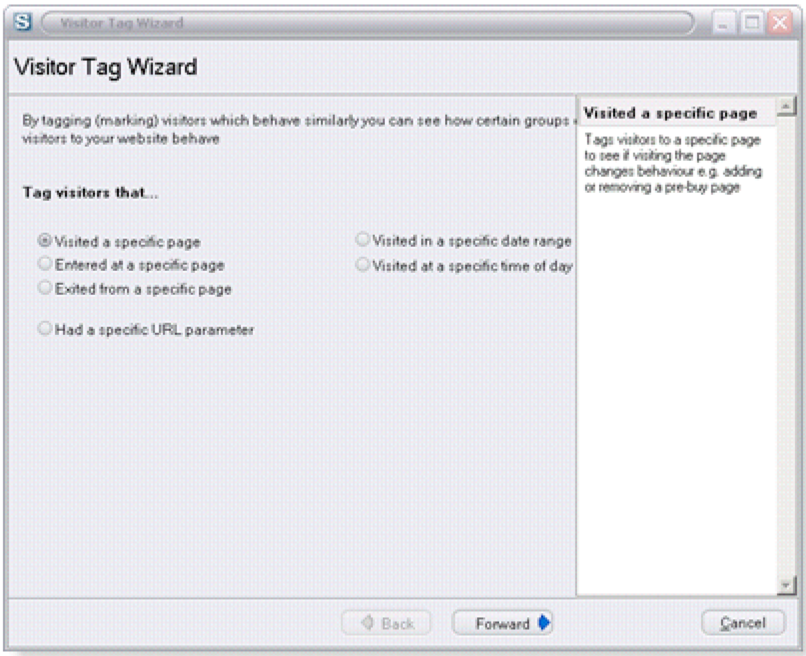
For each page defined a set of reports will be generated on the **Conversion** section when reports are generated for the site profile.

The **Revenue** column in the conversion tracking reports is obtained by multiplying the **Fixed amount** field value by the number of conversions.

Adding/Editing Visitor Tags

From the Setup Visitor Tags screen click the **Add** button to bring up the add or the **Edit** button to bring up the Visitor Tag Wizard. The options available will depend on the **edition** of SurfStats you are running.

The Conversion Page Wizard will display.



In the Professional and Enterprise Edition of SurfStats Website Traffic Analyzer the following options are available:

- Visited a specific page
- Entered at a specific page
- Exited from a specific page
- Visited in a specific date range
- Visited at a specific time of day.
- Had a specific URL parameter.

The wizard will take you through the steps involved in Adding or Editing a visitor tag. On the **Define** tab select or type a page name in the **Conversion Page** field. A list of page names will only be available after report generation. We suggest first generating reports to build a list of pages and then editing the visitor tags. You can also type the page name or part of the page name directly, for example:

If the string from the Top Pages Report is:
/ surfstats/ download.asp
You could enter the string as:
/surfstats/download.asp

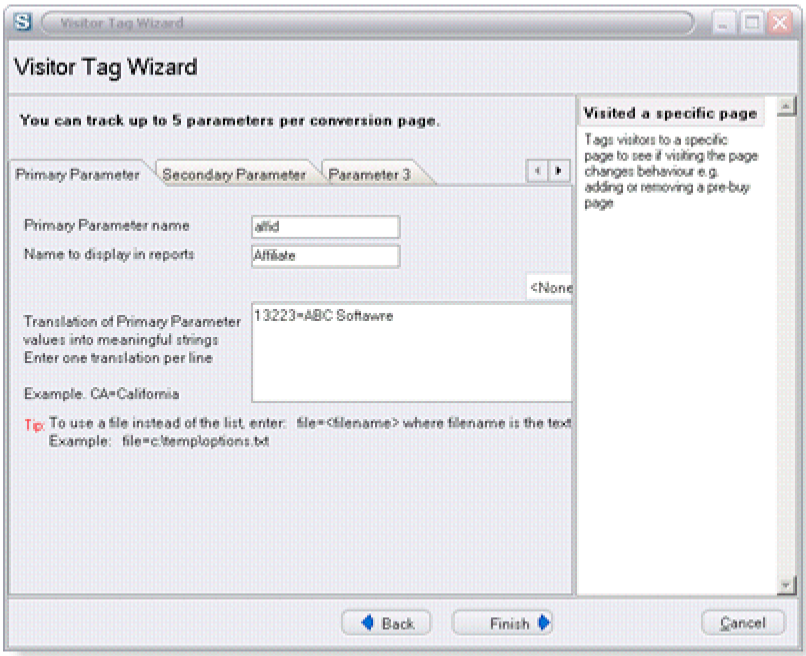
You can also use part of this string as wildcards for instance:
- /surfstats/
- download.asp
- download
- asp

Do not use wildcard characters like * and ?
(Use it only when defining a filter.)

Click the **Forward** button after the fields have been filled in.

URL Parameters

To perform URL parameter analysis select the **Had a specific URL parameter option**. For instance specifying a parameter as *affiliate* as shown below will create a list grouping the affiliates in the Conversion Tracking - URL Parameters report and will replace occurrences of *affiliate=13223* with *ABC Software* and occurrences of *affiliate=13224* with *XYZ Software*.



Specifying two or more URL parameters will result in matrix style reports generated in the URL **2D Parameter Analysis** and **URL 2D+ Parameter Analysis** reports.

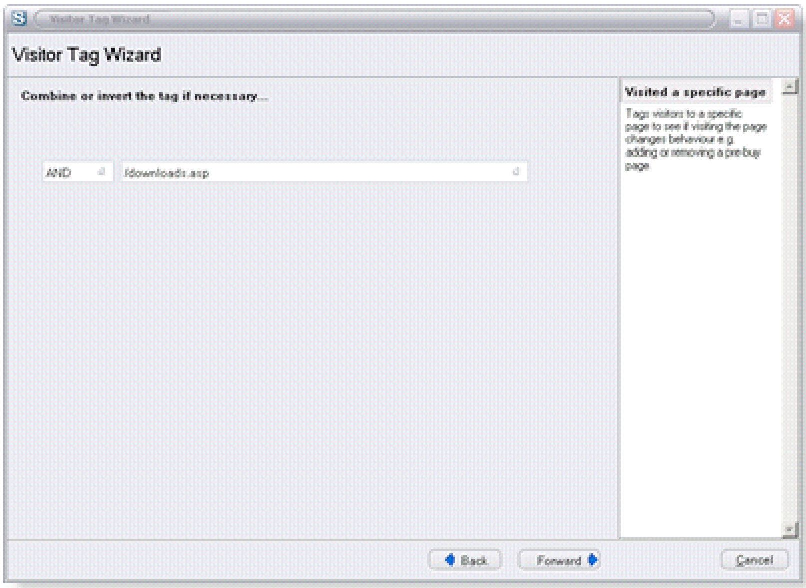
Click the **Finish** button after the fields have been filled in.

More URL Parameters

You can specify additional URL parameters on the Secondary Parameter, Parameter 3, Parameter 4 and Parameter 5 tabs. Click the **Forward** button after the fields have been filled in.

Combining Tags

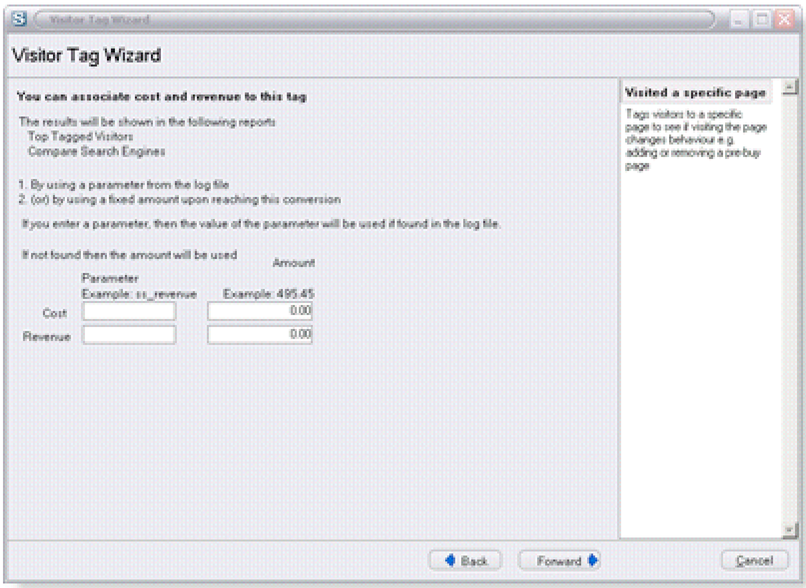
After specifying parameters click the Forward button to advance to the **combine or invert tag** step. Here you can narrow the selection by enforcing another restriction, for instance **From Search Engines**, which only looks at visitors that came from search engines.



Cost and Revenue

The next step in the **Visitor Tag Wizard** allows you to specify cost and revenue associated with reaching a conversion page. If you:

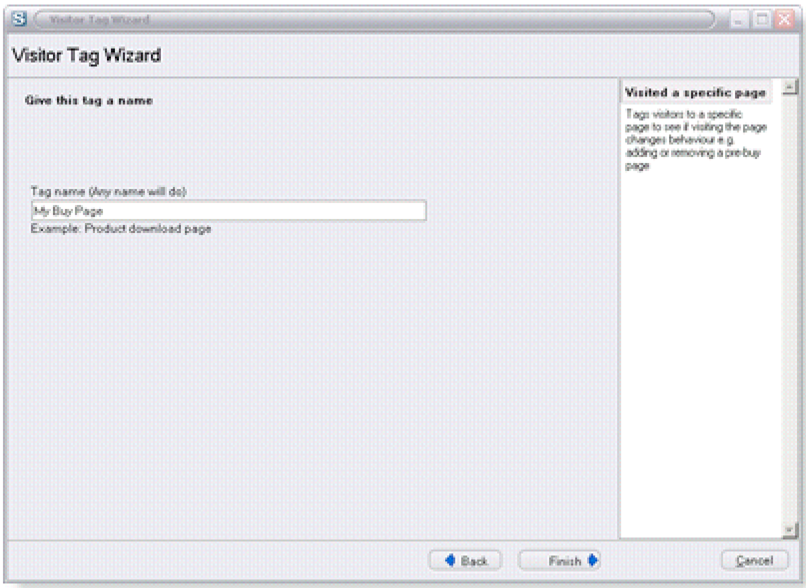
1. Specify a parameter the value of the parameter found in the log file will be used as the cost or revenue value. For instance if *ssrev* is specified as the **Revenue** parameter for the conversion page *buyloganalyzer.asp*, an entry found for *buyloganalyzer.asp?ssrev=95* will attribute a cost of $95 to reaching the page *buyloganalyzer.asp*.
2. Specify an amount for cost or revenue. This amount will be used as a fixed value when the conversion page is reached. For instance specifying a Revenue amount of 95 will ensure that $95 be attributed to the revenue field in the conversion tracking reports.



Click the **Forward** button after the fields have been filled in.

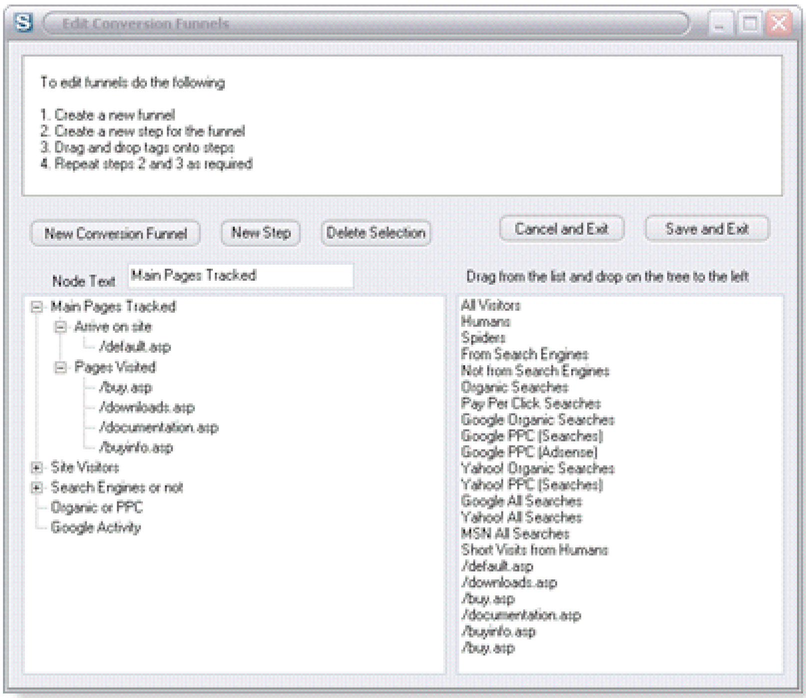
Visitor Tag Name

In the final step you can give the tag a friendly name. After this click Finish to complete the wizard.

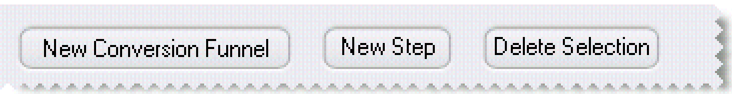


Conversion Funnel

You can define conversion funnels by clicking the **Edit Conversion Funnels** button while editing visitor tags (edit visitor tags by clicking the **Visitor Tagging** button at the top). The folowing screen appears:



A conversion funnel allows you to link more than one tagged page to a conversion funnel, which is for instance the steps in completing a shopping cart process. You need to specify a *Conversion funnel name*, which would be the same for all the pages in the funnel, the *Step Name* and the *Step Number*, which should increment as the user progresses through the pages, which represent steps of the funnel.



Click the **New Conversion Funnel** button to add a new funnel. This will add a top node at the left, representing a new funnel. You can edit the node text by selecting the funnel and editing the node text at the top.

Select a funnel and click the **New Step** button to add a step to the selected funnel. You can edit the node text by selecting the funnel and editing the node text at the top.

To link pages or a specific action to a step select the item from the list on the right and drag it onto the step on the left.

After you have created all the steps for the funnel and the necessary pages click the **Save and Exit** button. You will be prompted for regeneration of reports.

After report generation the funnel reports will be available in the Scenario Analysis reporting section.

Proxy Analysis

What is Proxy Analysis?

Proxy Analysis is the analyses of log files created by proxy servers and firewalls and is only available in SurfStats Enterprise Edition. Proxy Analysis focuses mostly on traffic generated by users in an Intranet accessing an external network like the Internet. The primary source of information used from the log files is the URLs (Uniform Resource Locator) requested by Intranet users. The URLs are scanned for key phrases by SurfStats, which places each URL in a specific category. User information on who requested which URL is also available in the proxy log files making it possible to categorize which users requested which category of information.

Proxy Analysis Profiles

The process to generate reports for proxy log files are the same as for website traffic log files. Create a site profile for the site, specifying the site profile type as Proxy. Specify the log files to process. Specify other options and then generate the reports.

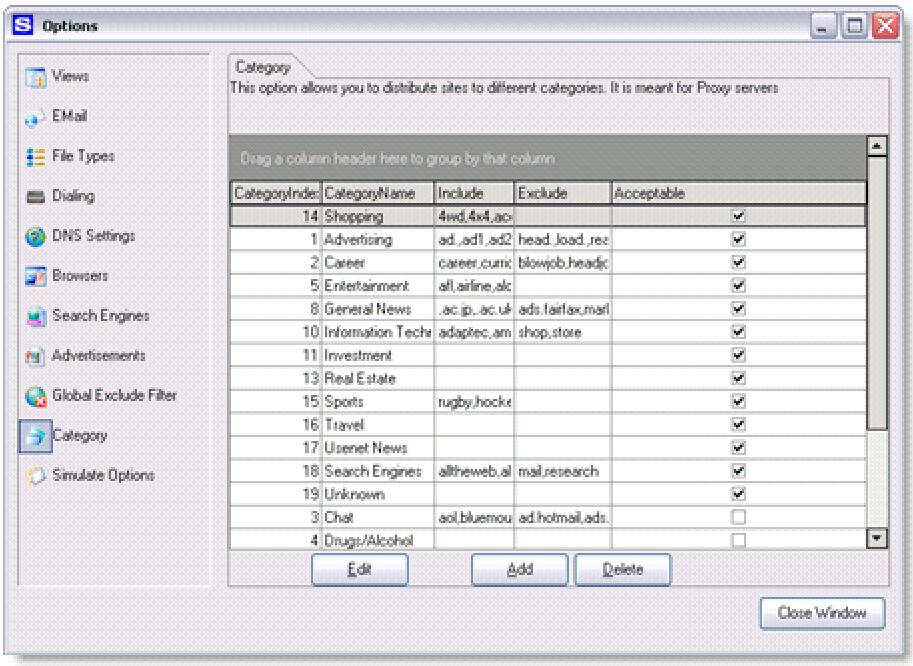
The **Create a new Profile wizard** is used to create a new profile for Proxy Analysis. Select a Profile Type of Proxy (instead of Traffic) to specify the site profile as a Proxy Analysis site.

Refer to the section **Adding a new site profile** on how to add a site profile. If you create a new profile from the Proxy Analysis tab the profile type is automatically set to Proxy.

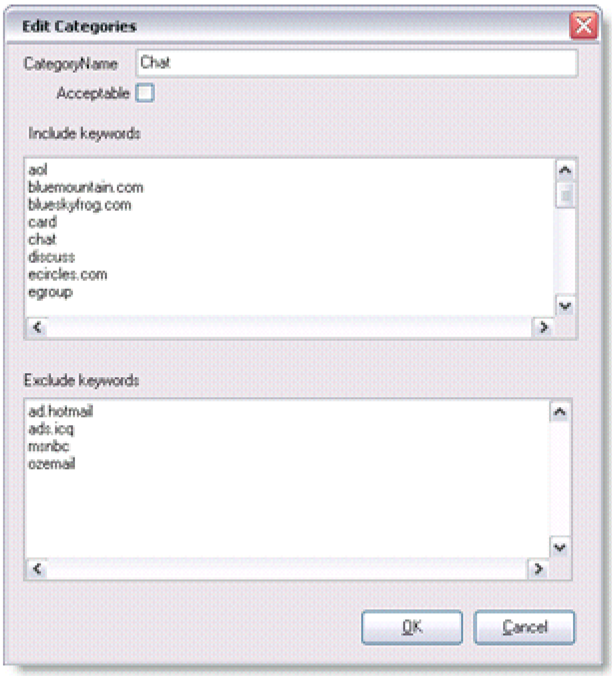
Category Setup

The phrases to be scanned for in the URLs can be edited (**Options - Category**).

Category setup links each phrase to a specific category of content accessed. Categories can be either classified as acceptable, like users looking for Investment information or unacceptable, like users looking for gambling sites, chat sites etc. You can decide what categories are acceptable and what categories are not acceptable based on the criteria applicable in your organization.



To edit a category double-click the category entry, which will bring up the following screen:



You can mark the category as acceptable or unacceptable by ticking or un-ticking the **Acceptable** checkbox. You can add keywords/search phrases to include or exclude by adding the keyword(s) at the end of the Include keywords or Exclude keywords list (type the keyword in one of the two scrolling list

The Include and Exclude keywords will be applied simultaneously to an URL to determine if the URL belongs to a specific category. For instance the URL http://www.card.com/msnbc will not be classified as belonging to the category **Chat** because although the include keyword card can be found in the URL, the exclude keyword msnbc is also present in the URL (assuming msnbc is listed as an exclude keyword).

Proxy Reports

Note that report names and what is displayed on the report category tabs (for instance Overview, Traffic, User etc.) is language dependent.

Overview Reports

Overview Reports provide a quick combined, graphical overview of the reports on the other report category tabs. The function of the overview reports is to get a general idea of site activity in a short period of time.

Traffic Reports

Summary Statistics – This report gives you an overview of general activity on the proxy server.

These reports give you an overview of site activity over a period of time. Refer to the Report Descriptions section under Site Traffic Reports for descriptions of the Hour of Day Activity,

Day of Week Activity, Day of Month Activity, Detail Daily Activity, Detail Monthly Activity, Views Per Visit and Length of Visit reports.

Detail Proxy Cache - This report gives information on what portion of requests were available in the proxy server’s cache and what portion of requests needed to be serviced by forwarding the request to another network like the Internet.

User Activity

Most Active Users – Displays information on the most active users (as grouped by IP address) in terms of page views, client sessions and bandwidth.

Authenticated Top Users – Displays information on authenticated top users (users that logged in with a user name and password) in terms of page views, client sessions and bandwidth.

Pages

The reports on the Pages tab display details about pages and files that were accessed or downloaded by visitors to your web site. Please refer to the Report Descriptions section under Site Traffic reports for more information.

Browsers

The reports on the Browsers tab display details about browsers used by site visitors. Please refer to the Report Descriptions section under Site Traffic reports for more information.
Sites Visited

Most Visited Sites – Displays information on the sites most visited by users in terms of page views, client sessions and bandwidth.

Top Acceptable Sites – Displays information on the sites most visited by users, which are categorized as acceptable.

Top Unacceptable Sites – Displays information on the sites most visited by users, which are categorized as unacceptable.

Sites visited by IP – Displays information on the sites most visited grouped by IP address in terms of hits, page views, visits and bandwidth. The sites are grouped per IP address.

Sites visited by Auth Users – Displays information on the sites most visited grouped by authenticated users in terms of hits, page views, visits and bandwidth. The sites are grouped per authenticated user.

URL Categories

Under URL Categories, the following reports are available:

Category Classification Activity – Displays information on URLs visited in terms of Acceptable and Unacceptable categories.

Category Summary – Displays information on URLs visited grouped by category

Most Popular Acceptable Web Pages – Displays information on web pages visited belonging to categories classified as acceptable.

Most Popular Unacceptable Web Pages – Displays information on web pages visited belonging to categories classified as unacceptable.

Errors

Refer to the Report Descriptions Section under Site Traffic Report for more information.

Complete

You can generate all proxy reports to folder by selecting File – Save Report to folder. Then use Windows Explorer to browse to the folder and double-click the file default.htm.

Streaming Media Analysis

What is Streaming Media Analysis?

Streaming Media Analysis is the analyses of log files created by Streaming Media Servers and is only available in SurfStats Enterprise Edition. Unlike a normal web server that downloads web content to a web client requesting a page or a file, a streaming media server streams multimedia content (audio and video) to clients. Content is normally delivered as unicast (content delivered to one client) or multicast (content delivered to multiple clients). The client uses a media player (like Windows Media Player for instance) to view and/or listen to multimedia content.

Creating Streaming Media Site Profiles

Streaming Media log files are analyzed the same way as web site log files are analyzed. Create a site profile for the site, specifying the site profile type as Media. Specify the log files to process. Specify other options and then generate the reports.

The Create a new Profile wizard is used to create a new profile for Streaming Media Analysis. Select a Profile Type of Media (in stead of Traffic or Proxy) to specify the site profile as a Streaming Media site. Refer to the section Adding a new site profile on how to add a site profile for Streaming Media Analysis. If you create a new profile from the Streaming Media tab the profile type is automatically set to Media.

Reports

Overview Reports

Overview Reports provide a quick combined, graphical overview of the reports on the other report category tabs. The function of the overview reports is to get a general idea of site activity in a short period of time.

Traffic Reports

The Traffic reports display site activity information for the period for which the reports are generated. This section of reports contains detail as well as summary reports.

Summary Statistics – This report gives you an overview of general activity on the proxy server.

The Hour of Day, Day of Week and Day of Month reports provides summarized media server activity information for the period being analyzed. This means for instance that if you analyze log files containing three months worth of data SurfStats will add up and group the data for hours 00-23, Monday to Sunday, day 01-31 respectively for the three types of reports.

The Detail Daily, Detail Weekly and Detail Monthly reports provides detail site activity information for the period being analyzed. This means for instance that if you analyze log files containing three months worth of data SurfStats will show report data for ninety days for the Detail Daily (or the limit imposed by the maximum allowable number of graph/table entries for the profile), twelve weeks for the Detail Weekly and three months for the Detail Monthly report.

Hour of Day Activity - This report shows the total hourly activity for each hour for the report period. This information is useful to determine for instance what time of day is best to do system maintenance. Note that the report is generated in the system time of the host server and then adjusted by the GMT Time offset (as specified on the General tab for the site profile). Statistics is shown for packets transferred, clip requests, client sessions and bandwidth used.

Day of Week Activity - This report shows the activity for each day of the week for the report period.

Day of Month Activity - This report shows the activity for each day of the month for the report period.

Detail Daily Activity - This report shows the activity for each day of the requested report period.

Detail Weekly Activity - This report shows the activity for each week of the requested report period.

Detail Monthly Activity - This report shows the activity for each month of the requested report period.

Detail Daily Activity - This report shows the activity for each day of each month for the requested report period.

Views per Visit - This report shows the number of clip views requested per client session. This report gives an indication of how compelling your site content is for visitors when they reach your web site. The more clips viewed the more compelling the media on the server.

Length of Visit - This report shows the duration (in minutes) of visits to your site. This information is an indication of how compelling your site is to visitors. The average US user spends less than one minute on a site. Strive to increase the time spent on site.

Visitors

Who is on – This report displays which users accessed the media site in the past 6 hours. If you are analyzing log files older than 6 hours no data will be displayed in this report.

Top Visitors - This report shows the IP address and/or the Whois/domain name and their relative activity level on the media server. SurfStats Log Analyzer does not differentiate between different users using the same IP Address. However it can group together visits from different IP addresses for a user with the same cookie (see section on Global Settings).

Top Visitor Organizations - This reports shows the most active organizations that visited your site. This is determined from the domain name of the visitor.
Tip - You must resolve IP Addresses or have the Visitor Origin Database installed before requesting this report.

Top Visitors by Bandwidth - This report shows the top visitors sorted by bandwidth.
Authenticated Top Visitors – This report shows detail activity of authenticated users accessing files and pages that require authentication (a login user name and password).

Top Visitor Countries - This reports shows the countries where your client’s domain names are registered. Please note that it may not be the actual location of your client. Some log analyzers assumes that .com and .net domains are located in the US but this is not always the case.

Top Visitor Regions - This reports shows the regions where your client’s domain names are registered. Please note that it may not be the actual location of your client.

Views per Visit – Shows the number of clips compared to the number of visits.

Length of Visit – Shows the number of length of time visitors spent on the server.

Pages

Top Audio Files – Displays summary and detail information on packets, % of total packets and bandwidth per audio file.

Top Video Files – Displays summary and detail information on packets, % of total packets and bandwidth per video file.

Top Unknown Files – Displays summary and detail information on packets, % of total packets and bandwidth per video file.

Most Downloaded File Types – Displays summary information on number of files and bandwidth per file type. The report is sorted by file type.

Most Accessed Directories – Displays summary and detail information on the number of times directories has been accessed.

Browsers

Refer to the Report Descriptions Section under Site Traffic Report for more information.
Capabilities

Refer to the Report Descriptions Section under Site Traffic Report for more information.
Watches

Refer to the Report Descriptions Section under Site Traffic Report for more information.
Referers

Refer to the Report Descriptions Section under Site Traffic Report for more information.
Errors

Refer to the Report Descriptions Section under Site Traffic Report for more information.

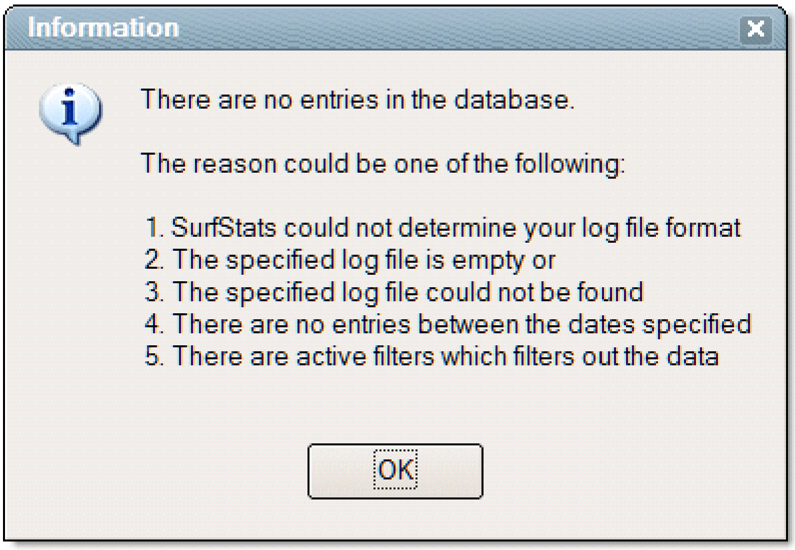
Complete

You can generate all streaming media reports to folder by selecting File – Save Report to folder. Then use Windows Explorer to browse to the folder and double-click the file default.htm.

Troubleshooting common errors

No entries in database message...

After SurfStats has processed the log files for a site profile you get the following information dialog:



You can go through the following checklist to help you to determine where the problem might be:

1. Check if the log files you are processing are valid. Check if you can generate reports from the first sample supplied with SurfStats.

   If you do not get reports from the sample - The application was not installed correctly or the sample was deleted. Remove the SurfStats application with Add/Remove programs and re-install.

   If you do get reports from the sample - Copy your log file to your hard disk and uncompress if it is an archive file (.zip, .gz etc.). Create a new site profile by dragging the log file (using Windows Explorer) onto the site profile panel in SurfStats. Double click on the site profile to generate reports using the log file. The program cannot process your log file if you get the same error as above. You can send a small zipped log file to SurfStats Support if it is a valid log file so we can try to determine what is causing the problem.

2. Check if the file specification (log file path and name(s)) is correct (on the Input tab of the profile). Note that you must specify the path and the file.

3. Check if the application has read rights to the log file by browsing to the log file(s) with Windows Explorer.

4. Check if you have the correct filter settings. The best way to check it is to have no settings at first. Then add one filter at a time.

A common error is that the program was installed on an older version before uninstalling the older version completely. Uninstall an older version AND delete the Surfstats directory before installing a new version.

Blank reports - without an info dialog...

If you get empty reports (No detail for period.... in reports message), without the 5-reason dialog box coming up:

1. Please see this issue first: http://www.surfstats.com/helpdesk/issue_view.asp?ID=311.

2. Ensure you have not installed SurfStats in the same folder as a previous version of SurfStats Website Traffic (Log) Analyzer.

3. Ensure you do not have duplicate versions of a file called C4DLL.DLL on the computer with different dates and sizes. Please rename them and replace with the one present in the SurfStats installation folder if you do find duplicate versions.

4. Ensure you have installed SurfStats while logged in as a member of the *Administrators* group.

A common error is that the program was installed on an older version before uninstalling the older version completely. Uninstall an older version AND delete the Surfstats directory before installing a new version.

No reports with the Scheduler...

If you can generate reports with the main SurfStats application, but not with the Scheduler, then the error is most likely caused by insufficient permissions or rights. When the main SurfStats application runs it runs under the user context that you are currently logged in as. When the Scheduler application runs it runs under the user context specified in Control Panel - Admin Tools - Services - SurfStats Scheduler Service - Log on tab.

1. If the log files are located or the output folder destination is on a different computer than where SurfStats is installed on, please ensure that you have changed the account under which the SurfStats Scheduler Service runs, from Control Panel - Administrative Tools - Services - SurfStats Scheduler - Log On tab. Do not use the LocalSystem account as it can only access drives on the local system. It cannot access any resources on the network.

2. Ensure that you are using UNC shares and not mapped drives when specifying log files on other computers. This is because mapped drives are only available when a user logs into the computer.

3. Ensure that the SurfStats Scheduler Service is running by checking the service status in Control Panel - Admin Tools - Services.

4. In SurfStats - Schedule Tasks - View Log have a look at the log for troubleshooting.

A common error is that the program was installed on an older version before uninstalling the older version completely. Uninstall an older version AND delete the Surfstats directory before installing a new version.

How to...

Adjust the time zone

Most log files are generated according the server system date and the time zone offset is usually indicated in a log line as in the following log line:
206.126.141.154 - - [10/Jan/2000:02:01:44 **-0500**] "HEAD /count4.htm HTTP/1.0" 200 0 "-" "-"
Some log files, e.g. MS IIS, are generated in Greenwich Mean Time (0 GMT).
It is recommended to always use the server system date.

In case you want to change the log file date, you can adjust the time zone for an individual site profile. To do so:
Select the site profile
Click on **Edit Profile**
Then on **General**
Then edit **GMT Time Offset**

Recommendation: For IIS web servers (which logs the time as GMT time) set the **GMT Time Offset** on the profile in SurfStats to what is specified in Control Panel – Regional Options. For instance if your time zone is GMT-07:00, enter -7. For Apache and most other web servers (that log the time as server time) specify the **GMT Time Offset** as 0.

Set up Mapped Drives

When generating reports manually, each profile has a default output e.g. (screen, email etc.) and you just need to be logged in. A mapped drive is attached when you log in.

When generating scheduled reports, one or more scheduled tasks can be created from a profile. Each of these scheduled tasks can be sent to a different output (screen, folder or email). These are run as a Windows Service. When running a report in a service, the default permission to be used is LocalSystem. This is good enough if you just want to read from the local computer and output to the local computer. However, if you want to place the results onto another computer in the network or read from another computer, you have to change the properties of the service because the service logs on into its own user context. In other words, it logs on as another user named LocalSystem by default and does not have any shares connected. It does not know about your mapped drives. You must specify these remote drives by using something like:

\\machine\surfstats\yesterday\access.log

where machine is the computer you want to connect to

or something like:

\\machine\C$\surfstats\yesterday\access.log

where C$ is the administrative share for the C drive on the computer called machine.
Note that if you specify log files to be analysed, the specification must contain the path and the file name/s (which can contain wild cards). If you specify a folder for output, the specification must contain the path and the user/scheduler must be able to create reports in this folder.

To set up the right properties for the Scheduler, go to Control Panel - Administrative Tools - Services and right click on SurfStats Scheduler. Then select Properties and click on the Log On tab. Make sure "This Account " is selected and it should contain the user that runs the software and has access to the remote drives.

Bandwidth detail in reports

Make sure that your web server or the web server at your host logs this information in the log files. It is set up as follows on IIS Servers: Open Internet Services Manager. Go to the web server entry (normally “Default Web Site”) on the left, right-click on it and click Properties. Go to the Web Site tab. Click the Properties button next to the log format list (normally it displays “W3C Extended Log File Format”). Click the Extended Properties tab. Make sure Bytes Sent (sc-bytes) and Bytes Received (cs-bytes) is checked. It is actually the easiest to check all the extended properties so that all extended properties are logged. Stop and Start the web server from Internet Services Manager afterwards.

Separate log files in reports

If you have a combined log file for A+B+C (all hits for site A,B and C are in one file) and you want separate reports for each site, you can do the following:

Set up a site profile for each site – one for A, one for B and one for C

Specify the same log files for every profile.

You then have to use filters to include log entries to A, B and for C. Setting the filters will depend on how the log entries to A,B or C can be separated, e.g. if it is in a separate directory (say \HomePageA), you should use the Files Like Include filter and specify the unique part of the directory in the Files Like Include filter field as *\HomepageA*. Note that it must have the wildcard * on both sides. Also note that filters are case-insensitive.

Please note that SurfStats has to parse the log files repeatedly for every site profile. It would be better to have separate log files for every site profile.

Combined server reports

If you have multiple sites (say site A, B and C) or clustered servers, each with it's own log files and you want a combined report on them, you can do the following:

- Set up a site profile (you can name it ABC).
- Specify all the log files for site A, B and C. You can specify multiple files and/or directories as well as different file formats.
- Use the first option of database population (overwrite all) when creating reports for A+B+C. (This option does not check for duplicate date/time entries).

Filter on Authenticated Name

There is no authenticated name filter in the control panel but you can look at the Authenticated Detail report, find out the IP Address of the user, and then filter on this IP Address.

Uninstall the product

Please refer to the Installation section.

Change default report output page

Edit a site profile and select the Output To tab. Change the Name field to the default home page of your web site, e.g. index.html or default.asp.

Generate reports in a single document

The MS Word automation server runs out of memory when creating a file with a large number of graphics. Therefore multiple reports in MS Word format cannot be generated with the program. Only single reports can be generated in Word format by clicking on the Word icon displayed at the top right hand corner when a report is displayed.

Change to output from Html to PDF to generate multiple reports in a single document. This format is faster to generate, uses less space and is supported on more platforms than MS Word.

Change IP Address ranges in table

Go to SurfStats Options – DNS Settings – DNS Ranges and enter the start (e.g. 24.64.0.0) and end (e.g. 24.71.255.255) and the company's name (e.g. Shawcable), click Add.
You should not specify wildcards in the start or end IP addresses of a DNS Range. Note that if you add a DNS Range (on the DNS Ranges tab) it is added permanently and it is only removed when you delete the DNS Range record

Generate reports unattended

Use the SurfStats Scheduler.

Create multiple scheduled tasks

Click Schedule Tasks from the SurfStats Website Traffic Analyzer application.

Click on Add task.

Select the site profiles that you want to schedule tasks for. You can use Ctrl and Shift with a mouse click to select multiple sites. Enter the date and time and schedule for the task(s).

Click on Post processing.

You can enter the directory destination with a combination of date macros and the site macro. Please note that the “root” directory (for instance c:\reports) must exist.

Click Go and then on Close.

The program will create the sub-directories and run the scheduled tasks according to your set up.

SurfStats and Terminal Services

Please note that SurfStats may not work properly with all remote administration software. The easiest way to find out is to try it with the remote administration software you are using.

Previous versions of SurfStats Website Traffic Analyzer (pre V7.0) had a lot of limitations for running SurfStats Website Traffic Analyzer and the SurfStats Scheduler application. V7.0 and later makes use of a different database as well as in-process DLL’s (instead of out-of- process DLL’s) for report administration, report generation and scheduling. This makes running SurfStats Website Traffic Analyzer from a Terminal Services Client virtually the same as running it from the Terminal Services server. The following needs to be mentioned:
· If two users try to edit the same site profile at the same time, edit the same scheduled tasks at the same time or try to generate reports at the same time from two Terminal Services clients, SurfStats will not allow it and give a warning message.

Scheduled tasks can be administered from Terminal Services clients or on the server from the Schedule Tasks menu option from within SurfStats Website Traffic Analyzer. Scheduled tasks can only run on the server, as the SurfStats Scheduler Service and SurfSchedRun application is responsible for the running of scheduled tasks. This is because only one instance of a specific Windows service (in this case the SurfStats Scheduler Service) can run at a time independent of how many Terminal Services clients are logged into a Terminal Services Server.

Get help on reports

To get more information about a specific report please refer to the help section at the bottom of each report.

Reverse DNS lookup from command line

In Windows you can look up the host name of an IP address from the command line. To do this open a DOS box with Start - Run - cmd<Enter>. In the DOS box type the following command:

nslookup -querytype=PTR <IPAddress><Enter>

for instance nslookup -querytype=PTR 66.228.118.58<Enter>

will display:

Your local server/router name and address

The Name of the IP Address owner

Creating hits from External Links in IIS Log Files

Visitor Clicks on external links (links to files not on the local server, like ads) do not generate hits in the local server's log file. These hits are logged in the log file of the remote server.
It is sometimes useful see the website traffic to external links and this can now be tracked by SurfStats.
To do this, you need to do the following:

1. Insert the following files in the root directory of your site

   ssclick.js
   externtracker.gif

These files are only supplied to licensed users of SurfStats Website Traffic Analyzer V8.2 and later and SurfStatsLive V3.2 and later.

2. Edit your webpage on which you want to track external hits and add the following line before the </body></html> tag

```
<script type='text/javascript' src='.\ssclick.js'></script>
```

3. Clicks to external links will then be logged and can now be tracked with SurfStats.
A visitor click to an external link (say www.shop.org/shop.htm) will create a hit line in the local log file like this:

```
2006-03-16 13:12:12 64.34.162.134 GET /externtracker.gif url=http%3A//www.shop.org/shop.htm 80 - 80.236.176.201 Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+5.1)
```

Note: Clicks on Google adwords cannot be tracked with this method at the time of writing.

Track 404 errors in IIS

You can customize your 404 File Not Found page in IIS 4 (and later) and make sure that the requested file is logged so you can correct the problem later.

Step 1
Create a custom 404 page that will redirect the user to a default page (say http://www.mysite.com/default.asp) and name it 404.asp or 404.aspx
It must be an asp or aspx file so the IIS server can write the information to the log file. Put the following in this active server page (Change mysite to your domain):

```
<% Response.AppendToLog "_" %>

<html>
<head>
<meta http-equiv="refresh" content="0;url=http://www.mysite.com/default.asp">
<title>Error 404 at www.mysite.com</title>
</head>
<body bgcolor="#ffffff">
<p> </p>
</body>
</html>
```

Step 2
Place this 404.asp file in the root directory of your website.

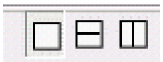
Step 3
Now configure IIS to use this 404 page (/404.asp) when a file is not found..
Start - Programs - Administrative Tools - Internet Information Services (IIS) Manager. Select your website, right click on it and select properties. Select 404 and click on Edit. Choose message type as URL and enter /404.asp in the URL field.

When a 404 error is encountered, IIS will direct the user to the custom 404.asp page which will create the entry in the log file and redirect the user to your default page. These entries can then be tracked with SurfStats (as from V8.2 SurfStats Website Traffic Analyzer and V3.2 SurfStatsLive).

Create a Date Split

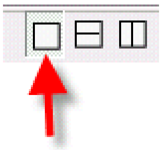
A Date Split is used to compare Website traffic between two different periods, e.g.last week's file requests to this weeks file requests.

To create a date split, click on one of the split buttons when a report is displayed.



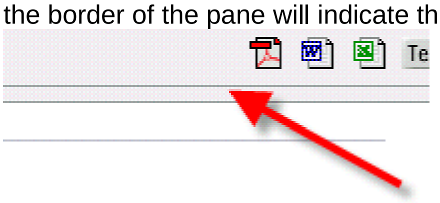
The second split button creates a horizontal split and the third split button creates a vertical split.

You can remove the split by clicking on the one page button.



To select a date range for the splits, click on "Calendar" and then on "Date Split". Select a date for the split

To select the active pane after the split, click on the thin bar at the top of the pane. A blue line around the border of the pane will indicate the active pane



Get more than 20 results in a report

To get more than 20 results in a report:

- Edit the site profile and click on Report Views (under General).
- Open the Page Reports (report views panel) and click on the Top Pages.
- Set the number of Main Table Entries and click on the button next to the Main Table Entries.
- Reparse the log files.

Multiple reports from one log file

Some ISP's allows you to sub-host other sites on a virtual server. The site content (web pages and other files) is stored in a separate virtual directory (and subdirectories of it) for each site hosted on the virtual server. The site activity for all sites is logged to one log file (per day, per week etc.) on the virtual server. For instance you can have an account with an ISP and host sites for www.john.com and www.paul.com. The site content for www.john.com gets stored in the virtual directory /john and subdirectories of it. The site content for www.paul.com gets stored in the virtual directory /paul and subdirectories of it. Activity on these sites gets logged to a single log file (per day, week etc.) located in for instance c:\winnt\system32\logfiles\W3SVC48, which is the physical path mapped to for instance the virtual directory /root/logfiles/W3SVC48.

The approach followed here is to create one profile that will retrieve the log file(s) from the FTP server on the virtual server and store the files locally on the computer running SurfStats. A profile is then created in SurfStats for each site hosted on the virtual server to analyze the log files (locally) that has been retrieved by the FTP profile.

Profile to retrieve log files

The first step is to set up a profile to retrieve the log file(s) from the server with FTP. This profile only retrieves the log files without doing any reporting. You only need to set up one of these profiles even if you host more than one site on the virtual server. For this step you need to decide on the period you want to report on and set the appropriate date macros to retrieve the correct log file(s).

Click the New Site Profile menu item to start up the New Profile Wizard. Give the profile a name like GetLogFiles for instance and set the profile type as AutoFTP. Click Forward to advance to the next screen of the wizard.

On the next step of the wizard fill in the FTP login details to retrieve the log files. For the Default Folder specify the folder where your log files for the specific virtual server are located. Click Forward to advance to the next screen.

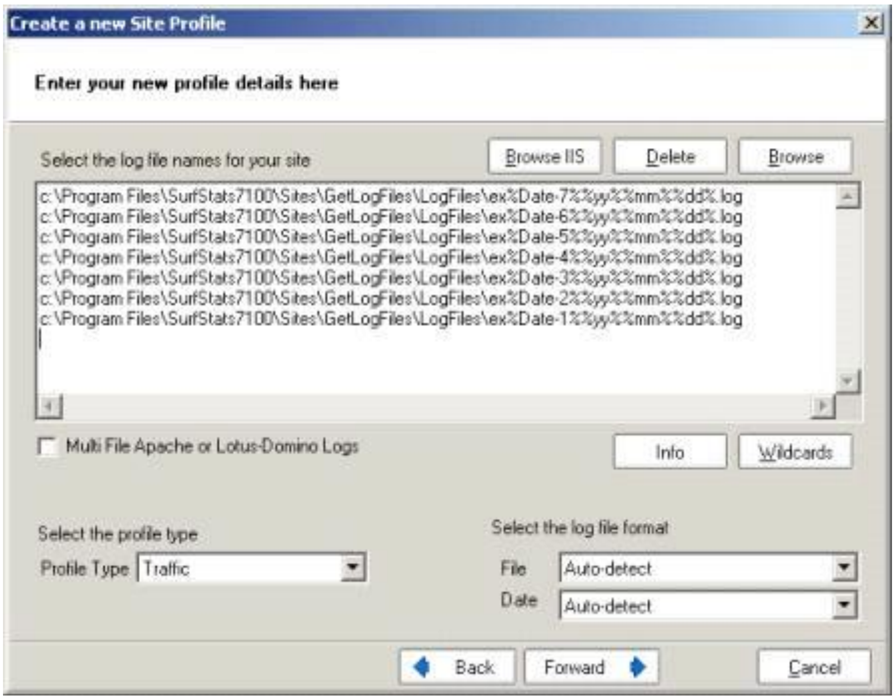
The next step of the wizard is where you specify the name(s) of the log file(s) to retrieve. Complete the rest of the steps of the wizard. Select to Return Control to Main Screen as the action to perform after processing log files. Click Finish afterwards to complete the wizard.

Set up a profile per site

Set up a profile per site

The next step is to set up a profile to generate reports for a site from the log files retrieved with the site profile in the previous step. Set up a site profile (with the New Profile Wizard) for each site hosted on the virtual server (profile of type “File”). For example set up a site profile called john for www.john.com and a site profile called paul for www.paul.com.

Click Forward to continue. The next screen shows how the log files are specified for the john profile (it would be exactly the same for the paul profile). This profile will analyze the log files from the past 7 days located in the folder where the FTP profile dropped the files that it downloaded (in the previous step).



Click Forward and complete the rest of the steps and Finish to complete the wizard. You would probably want to generate the reports to a folder. Create profiles in a similar way for each site.

Click on each of the newly created profiles in this step (john and paul for instance) and click Edit Site Profile. Click the Filters menu item and specify the Files like include filter as the virtual directory where the site content is located for the site on the virtual server. For instance for the www.john.com site the site content is located in the /john virtual directory so specify *john* as the Files like include filter on the john profile, for the www.paul.com site the site content is located in the /paul virtual directory so specify *paul* as the Files like include filter on the paul profile. Click OK to save changes to the profile.

Generate reports per Profile

To generate reports first of all make sure that the SurfStats dialing options is set up correctly if you do not have a permanent connection to the Internet (SurfStats Options - Dialing) so that SurfStats could initiate a dial-up connection to the Internet if required.

Double-click on the site profile (GetLogFiles for instance) created in the first step to retrieve the log file(s) from the FTP server. Click Start and wait for the profile to finish downloading the log files from the FTP server. You have to run this profile only once to obtain the log files.

After the log file(s) has/have been downloaded, double-click the first site profile (john for instance) that points to the downloaded log files (from the previous step) and generate reports for the profile. Do this for all the other profiles (paul for instance).

Appendix

Report accuracy

There are many factors that influence the accuracy of reports from analyzing log files. If the generated reports look inconsistent, check that:

The log files are parsed and entered in the database correctly. You can do this by specifying a small log file and comparing the reports to the actual entries in your log file.
You have specified the log files correctly.
You do not have many error lines in your log files that are being discarded by the program.
You have specified filters correctly. Note that filters are not case sensitive.
You have specified the period to report on correctly.
You have used the correct method to populate the database. We recommend that the first option (overwrite the database) be used to populate the database. Changing filters when using the Add to database option would mean that the database could contain inconsistent entries.

Unknown referrers

The report will sometimes show up some fields as "Direct Traffic" if the log line does not contain all the referrer information. The fields will show as "-" in the log line. This may happen when:

The user typed the URL of the web site directly into the Address field of the browser, i.e. the URL was not referred from another site.

Software was used that blocks the sending of browser information to the server.
The browser software does not send referrer information, e.g. some spiders and e-mail siphoning software.

Assumptions

Log Analyzers have to make certain assumptions when analyzing log files. These include how to track unique visitors (using IP Addresses or Cookies), the cut-off time to count a visitor as a new visitor and how errors in log files are handled during the parsing process.

A user session includes a series of consecutive requests from a user to an Internet site. A user session is terminated when a user does not make another request for more than 30 minutes. This is referred to as the cut-off time. This means a user may log in many times per day and if the requests are always less than 30 minutes in-between, the program will only report it as one user session.

The period being analysed also plays a role. If you determine the number of unique users every day and add it up to get a week's unique users, this would probably not be the same as the number of unique users determined per week. The reason for this is that a unique user might not be counted as a unique user early the next day when using the IP Address as unique identifier and the cut-off time is less than 30 minutes. A daily count would show 7 but a weekly count would show 1 when cookies are used and the same user visits every day.

Dynamic IP Allocation

Due to the shortage of IP Addresses, most ISP's dynamically allocate IP Addresses to their users from their limited range of IP Addresses. This means that a visitor with a certain IP Address might not be the same visitor with the same IP Address a little later.

Mapping individual IP addresses or requests onto users or even using cookies are not entirely accurate -- the exact number of individual users can only be determined by requiring them to log on individually to your site.

Caching

One problem in providing accurate results is the caching performed by your PC and ISP. Caching reduces the number of pages requested at the host server, as popular pages are stored locally either as temporary files on your hard drive or in the ISP's page cache. Since log files only record files requested from the server, the number registered in the log file can be incorrectly low.

Visitor's Disk Cache

If your web site has been visited, the downloaded files are on the visitor's hard disk. Requests for files on your web site would be first served from the visitor's cache and if not available (or old), would only then be served from your web site's server. This would result in fewer hits in your web site's log files.

There are JavaScript code available on the Internet that you can put in your web pages to force a reload or refresh of your web page. You can search for "force reload JavaScript" on a Search Engine to find a suitable script.

Caching Servers

Many ISPs are running cache servers to reduce web traffic and speed up the transfer of commonly accessed web sites by their users. There are also several national and international caching servers that many ISP's are using in order to speed up transfers. This means for instance that if one AOL user visits a site, then any of the other AOL users who visit that site after that would load the site from the AOL cache, and not from the visited site. The larger ISP's like AOL, Earthlink, MSN, Prodigy, MSN, Earthlink, etc., all use caching servers and hits from these users could be much more than reflected in the log file.

Geographical location

Visitors' locations cannot always be accurately determined as an IP address does not necessarily specify the geographical location of the user. A large number of IP addresses can also remain unresolved.

Hits from "non-visitors"
Spiders or robots from search engines that catalogue your site can exaggerate your site statistics, as they are not true visitors. Eliminating them from your overall scores can provide more realistic figures.

Firewall or proxy server
The incoming web traffic that passes through a firewall or proxy server to reach the web server is in most cases identified by a single IP address of the firewall or proxy server. This is called Address Translation. The log file will in a case like this only contain one unique IP Address even if multiple users from behind the same firewall or proxy server visited the web site. Analyzers will these visits as visits from the same user.

Errors in log files
Errors in log files are usually the first item to look for when you are getting inconsistent reports or reports that are not making sense. Errors in a log line may cause that the log analyzer discard the line's information. If your log files are not very large, you can edit the log file with an editor like Wordpad or Notepad. Look for strange characters or misplaced text strings and remove or correct it.

You should correct the cause of the problem (e.g. an overloaded server) if your log file contains many errors. Your reports will only be as accurate as your source data.

Page view and user sessions

Some clients have reported that page views and user sessions do not tie up. Please keep the following in mind when interpreting report results:

Files not defined as a page type and graphic files also trigger a user session.
The user session cut-off time.
Dynamic IP Address location.
The effect of caching.

One visit can include many hits. A report on downloaded files may show a few hits but no user session. This is because the visit is counted at the start page where a visitor enters the site. Only a request for the entry page will count as the start of a user session.

Tips & Tricks

General tips
Insist on referrer information in log files or find a host that provides log files with this information. You cannot get meaningful information of your web site's activity without referrer information.
If you have multiple sites, set up a site profile for each site and also a profile that combines all of them.

Use a third party submission service like World Submit to submit your site to search engines. When using a third party services your e-mail address is not supplied to search engines and you would not get the usual, mostly unwanted, e-mail replies.

Do not over analyze.

Adhere to the Parento principle (also know as the 20 - 80 percent principle) when analyzing a web site. This means that you should first focus on the few clients that generate most of the traffic to your web site, the few referrers that refer most traffic to your web site, the few countries that most of your clients come from, the few incorrect links that causes most 404 errors, the search engines that generates most of the traffic, etc.

Use keywords to improve popularity

Use the Top Search Keywords and Top Search Phrases reports under the Referrer heading to see which keywords or phrases visitors use to find your site. Gather statistics for a couple of months before you decide on your search phrases. Build focused pages around queries and ensure that you have pages within your web site with good, solid content for these terms to please the crawlers. Do not pick keywords that are too popular or broad. Once you have set everything up, add the keywords or phrases in the Meta tags and submit to search engines.

Look for visitors that never click past your homepage

If the number of single access pages is unreasonably high or visitors seldom click past your homepage, this can indicate a weak page that does not inspire visitors to further explorer your web site. An image that takes a long time to load on a main entry page and that is used, as an entry page is a definite No-No. When using a Flash entry page, also create a link to bypass the Flash entry page. The result will surprise you when you compare hits between the Flash and Bypass pages..

Look for visitors that bypass your payoff page

If a substantial number of visitors exit your site without visiting your payoff page, the pages linking to

your payoff page could be weak and may need some atention. Make sure you have clear links and a reason for visitors to visit your payoff page.
Look for incomplete hits by noting the bytes transferred
Incomplete hits can indicate server capacity and connectivity problems or that a page is too large. See if there are date/time patterns, which also indicate capacity problems.

Look for Quality, not just Quantity

Track hits from ads and reciprocal links by tracing their clicks through to your payoff page. It doesn't pay to run ads that generate hits but seldom click-through to your payoff page. Also see the section on advertising analysis.
Look for visitors repeatedly entering your site on pages other than your homepage
This could indicate the other pages have important search engine keywords in their text that should be included in your homepage text, or that you should add their Meta tags to your homepage Meta tags.

Look for patterns of search engine keywords that result in hits on your payoff page

This helps target the search engines, understand your visitors' needs, and select the HTML text that best motivates your visitors.
Look for referrers that generate a lot of traffic to your site
Visit the site that generate a lot of traffic to your site and see why they linked to your site. This can help you to market your site more effectively. No referrers from search engines indicate that you are either not listed or not listed effectively on search engines. You can check your site's popularity on the web with services that provide link popularity checking like www.LinkPopularity.com They provide this service for free and offer tips on how to improve your link popularity.

Look for search robot activity

Add a robots.txt to the root directory of your web site. Search robots can generate a lot of hits, especially to the robots.txt file. Exclude (filter out) accesses to the robots.txt file by search engine robots as this can influence or distort traffic significantly. If you get no requests for this file it means that you should submit your site to search engines. No hits to the robots.txt file indicates that your are not listed on search engines that send out robots to catalog your site.

Look where clients exit your web site

This can be the last page they accessed or they can exit it by clicking on an interesting link on your page (also refer to the section how to create hits to external links in your log files in Advertising Analysis).

Look for web site problems

Web site problems that completely break a web page are pretty easy to spot (like very few hits or error codes in your log file), but others may be more difficult to find. If your site gets a large number of unique visitors, but the total number of page views is small, the problem may be due to:
Poor site navigation
Uninteresting site content
Browsers that cannot view your site’s pages properly
Use the detail daily error report and the other error reports to monitor, find and reduce errors on your web pages. Start with the most errors of incorrect links and fix it.

Do visitors find your site interesting?
See if clients bookmark it. Add a "favicon.ico" file to the root directory of your web site. Also add the "favicon.ico" file to every directory of your site. When a visitor, using Internet Explorer 5+, bookmarks your site, a request is made for the file "favicon.ico" You can then filter on the file "favicon.ico" to see the number of hits. By knowing the number of hits from the total browser hits and those just Internet Explorer 5+, you can get a good ballpark figure of the total number of bookmarks to your site. By generating a report on hits to the "favicon.ico" file over time, you can see if you are getting new visitors and if the number is improving. (Repeat visitors have probably book-marked you page already and will not bookmark it again). By placing a "favicon.ico" file in every directory, you can also see which directories are bookmarked.

See if there are many single access pages in relation to the number of hits per page. The alarm bells should ring if the number of single accesses to default.htm (your home page) is about the same as the

total number of accesses to this page.
See the number of accesses to your home page (usually default.htm) in relation to accesses to all other pages. This is a good indication of how many visitors click past your home page.
See if visitors recommend your site to their friends and associates. At http://www.recommend- it.com/ you can get a script that visitors can click on to recommend your site and also informs you when it has been recommended.
See if they link to your site. You can find out by typing the following in the search window of most search engines:
Link: www.yoursite.com where yoursite.com is you web page address.
See the time visitors spend looking at your pages.
See how many return visitors you get.
See the number of page views per visitor.

Find out more about visitors
If you have received e-mail from a visitor, you can look at the suffix of his e-mail address and then look-up the country from the ISO country code list. This list is available in the online help in SurfStats (Help-Contents-Appendix-List of country and ISO internet codes).
You can also do a Domain Name System (DNS) lookup of a visitor's IP Address, also referred to as IP Address Resolution. A DNS lookup queries DNS servers for the IP Address of a domain name. Note that some IP addresses cannot be resolved due to a variety of reasons like some ISP's allocating dynamic IP addresses to their clients, the host is not available, etc. The resulting URL may contain the domain name and the country of origin of your visitor.
You can also find out more about your visitors by clicking on the IP Address hyperlink in a report and browsing to your visitor's web site.
You can query IP Address databases through a whois lookup at the three Regional Internet Registries (RIRs) worldwide, which collectively provide IP registration services to all regions around the globe.
ARIN (At http://www.arin.net/whois/index.html) - North America, South America, the Caribbean and sub-Saharan Africa.
RIPE (At http://www.ripe.net/db/whois.html) - Europe, Middle East, parts of Africa
APNIC (At http://whois.apnic.net/apnic-bin/whois.pl) - Asia Pacific
or at
AllWhois at http://www.allwhois.com/
Network Solutions at http://www.networksolutions.com/cgi-bin/whois/whois
Department of Defense athttp://nic.mil/cgi-bin/whois
GeekTools at http://www.geektools.com/
UXN Spam Combat at http://combat.uxn.com/

Look for server attacks
Server attacks are worm or virus attacks as well as unauthorized users trying to gain access to protected directories.
A common server attack is a request to the "cmd.exe" file caused by the Nimda or the Code Red virus present on a Windows machine. If the machine is on a local network, you can find out which machine it is on and inform the network administrator. If this file was requested from the Internet, you can do a DNS lookup through the program or do a Whois lookup (usingwww.allwhois.comor www.arin.net) to find out who this comes from. If you can find out where the request came from, you can inform the originator of the virus on their server.
These requests usually cause many 404 (File Not Found) errors in your log files. If you find successful accesses to the file cmd.exe on your Windows Server, your server is probably infected.

Clean up your log files before processing
You could speed up processing of your log files if you remove unwanted entries (e.g. requests for images files, 404 errors, junk entries, etc.) from your log files first. You can do this with the free SurfStats FileFilter available at:
http://www.surfstats.com/download.asp

Glossary of terms

Absolute URL - The full Internet address of a page or other World Wide Web resource. The absolute URL includes a protocol, such as "http," network location, and optional path and file name. For example, http://surfstats.com/ is an absolute URL. See also URL.

Active Server Page (ASP) - A document with a .ASP extension that contains embedded server-side scripting. ASP-compatible Web servers can execute these scripts. On the client side, an ASP is a standard HTML document that can be viewed on any platform using any Web browser.

Ad - A text, graphic or banner advertisement on a web page that takes the visitor to another site when clicked on.

Ad Click - A mouse click on an advertisement on a web site, which takes the visitor to another site.

Authentication - The verification of a user by matching a username and password in a multi- user or network environment. A user's name and password are compared against an authorized list, and, if the system detects a match, access is granted to the extent specified in the permission list for that user.

Bandwidth - The measure of the traffic (in kilobytes of data) transmitted from the web site.

Bookmark - Also referred to as an anchor. This is a named location on a Web page that can be the target of a hyperlink. A bookmark can be applied to a string of characters or exist on a page separately from any text. Bookmarks allow authors to link to a specific section of a target page. A bookmark is preceded by the hash sign (#) in an URL,.

Browser - The program used to locate and view HTML documents.

Cache – An area that stores text and image files from a web site already viewed. When returning to the web page again, the browser loads the files from the cache, thus saving time.

CGI (**Common Gateway Interface**) - A standard method of extending Web server functionality by executing programs or scripts on a Web server in response to Web browser requests. A common use of CGI is in form processing, where the Web browser sends form data to a CGI script on the server, the script integrates the data with a database, and sends back the results as a Web page. Use of CGI can make a Web page much more dynamic and add interactivity for the user.

Client (visitor or user) - A computer that accesses shared network resources provided by another computer on a local area network or the Internet. A unique IP address or URL identifies this visitor.

Client Session or visit - A client session or visit is started when a request for a file (of any type) was made from a unique IP Address. If no more file requests were made in the cut-off or idle time, the program assumes that the visit was terminated. If a file quest is then made after the cut-off time from the same IP Address, a new visit is counted. The program's default cut-off time is thirty minutes.

Client or visit tracking
There are three commonly used methods to track visitors.
1. The most accurate way to track unique visitors is by forcing them to log on every time they visit your site. This way the authenticated name is used to track a client's activity on the site.
2. Using cookies can improve the accuracy of tracking unique visitors. The cookie (saved on the client's pc) is used to differentiate between visitors that used the same IP Address.
3. Tracking file requests from IP Addresses. Determining the uniqueness of a visitor by tracking their IP address is not always accurate because most ISP's use dynamic IP Address allocation. The following illustrates this:
· A visit can be incorrectly counted as a new visit. The same user can have a different IP Address if the session was terminated and he logged on again. The program will count this visit as a new visit.
· A visit can be incorrectly continued. If the IP Address was allocated to another user, which logs on to your site within the cut-off time, he will be seen as the previous visitor.
· Some ISPs provides the same client with different IP Addresses during a logon session. This can increase the number of unique visitors significantly.

Client error codes - Status Error Codes in the 4XX range indicating the type of errors that occurred during a file/page request. Also refer to the Status Codes.

Cookies - Files containing information about visits to a web site (e.g., user name and preferences). The user's browser provides this information during the visit to a Web server. The server informs the browser to store some of this information (and perhaps other) on the client machine in a cookie file. When the visitor accesses the same web site again, the cookie is automatically presented to the server by the browser and the server can then configure it based on the information provided or react upon the information in the cookie.

Database - A file containing records of information that are organized and presented to serve a specific purpose, such as the facilitation of searching, sorting, and recombination of data. Databases can be published on the World Wide Web to let site visitors look up information in records or add new information to the database.

Directory - World Wide Web pages consist of varying files. These files are often set up in a directory structure, resembling "tree branches". Directories are usually structured to keep similar topics within a specifically named directory.

Domain Name - The text name corresponding to the numeric IP address of a computer or domain on the Internet.

Domain Name Lookup - Also called DNS lookup or IP Resolving. This is the process of converting a numeric IP address into a domain name or URL.

Domain Name Suffix - The last digits of a domain name can be used to identify the country or type of organization. Possible suffixes for the organization type includes:
.com = Commercial
.edu = Educational
.int = International
.gov = Government
.mil = Military
.net = Network
.org = Organization
.xx = where the xx is a two digit country code, e.g. .uk for United Kingdom

Download Manager – A program that is used to download large files from a host computer. This type of program can usually resume the download process after an aborted file transfer thus reducing bandwidth requirements. There are various download managers available like GetRight, Gozilla, ReGet, BullitproofFTP, NetVampire, etc.

Entry pages - This is the first requested page from a client or how the client enters your site. If the client entered your site from a search engine, this would be the file name part of the URL submitted to the search engine. The server looks for the default page that can be any of the following in the order:
default.htm
index.phtml
index.pht
index.cgi
index.shtml
index.html
welcome.cgi
welcome.html
index.htm
index.php3
The first page that it finds will be displayed and if none of these files are found, it will display the list of files in the directory.

Exit pages - This is the last page your client has requested.

External hyperlink - A hyperlink pointing to a page or file that is outside of the current web, i.e. not on the current server.

Favicon.ico - An icon file that is placed on a visitor's links to a book-marked site (only in Internet Explorer 5 or later). Hits to this file can be an indication of how many visitors bookmarks your site. Also see http://www.favicon.com/ how to generate one.

Filters - A means of narrowing the scope of a report by specifying ranges or types of data to included or excluded.

Firewall - A method of protecting the files and programs on one network from users on another network. A firewall blocks unwanted access to a protected network, while giving the protected network access to networks outside of the firewall. A company will typically install a firewall to give users access to the Internet while protecting their internal information

Forms - An HTML page, which passes variables back to the server. These pages are used to gather information from users.

Folder - A named storage area on a computer containing files and other folders.

Frame - An area of a Web browser window defined by a frames page. A frame appears in a Web browser as one of a number of different areas in which pages can be displayed. A frame may be scrollable and resizable, and may have a border. You display a page in a frame by creating a hyperlink to the page and specifying the frame as part of the hyperlink.

Frames page - A page that divides a Web browser's window into different areas called frames that can independently display several Web pages.

FTP - File Transfer Protocol is a standard method of transferring files between computers over the Internet. You can do so with a FTP client program.

GIF - Graphics Interchange Format is a raster based image file format commonly used in HTML documents. The PNG format is slowly replacing this format as the GIF format is a proprietary format from Unisys.

Hit - A request for any object or file that is on a web site. This could be an html page, a file or a graphic on a page. A request for a page can generate a lot of hits depending on how many sub-elements of files the page consists of. This is an indicator of web site traffic but not an indicator of how pages were looked at. Also see Page and User.

Home Page - The main or default page in HTML (Hyper Text Markup Language) format of a Web site. This page usually provides visitors with an overview and links to the rest of the site. This page's address is referred to as URL.

HTML - Hyper Text Markup Language is used to write documents for the Internet. It specifies hypertext links between related objects and documents.

HTTP - Hyper Text Transfer Protocol is a standard method of transferring data between a Web server and a Web client.

IP Address - The Internet Protocol address identifies a computer connected to the Internet.

Log File - A file created by a web or proxy server, which contains all of the access information regarding the activity on that server.

META tag - An HTML tag that must appear in the HEAD portion of the page. META tags supply information about a page but do not affect its appearance. These tags are used by some search engines when adding a web site to the catalog.

Organization - A commercial, academic, non-profit, government, or military entity that connects users to the Internet, identified by an entity's Internet domains.

Page - Any document, dynamic page, or form defined as a page file, e.g. .htm pages..

Page (HTML) View or Request - The request for a file defined as a page file. The page is basically what you see after the transfer and can consist of many other files. Page requests do not include hits to images, component pages of a frame or other non-html files. The number of page Hits = the number of page views. Also see hit.

Platform - The operating system (i.e. Windows 95/98, Windows NT/2000, Linux, Unix, Macintosh, etc.) used by a visitor to the site.

Protocol - An established method of exchanging data over the Internet (e.g. FTP, HTTP, etc).

Proxy server - An Internet server that acts as a firewall, mediating traffic between a protected network and the Internet.

Referrer - The URL of an HTML page that refers to the site. Referrer data is useful to determine how users came to your site.

Relative URL - The Internet address of a page or other World Wide Web resource relative to the Internet address of the current page. A relative URL gives the path from the current page to the destination page or resource. A relative URL can include a protocol. For example, the relative URL Test/Sample.htm refers to the page Sample.htm in the folder Test, below the current folder.

Robots - Also called search spiders. These are the programs used by search engines to catalog web sites.

Robots.txt - A file that instructs the search engine's robot which directories to catalog or to skip. If you do not have this file, it will show up as a 404 error (file not found) when a search robot visits your site. You can use an empty file.

Root web - The default, top-level web provided by a Web server. To access the root web, you supply the URL of the server without specifying a page name or sub-web.

Script - A type of computer code that can be directly executed by a program that understands the language in which the script is written. Scripts do not need to be compiled to be run.

Status Codes - A code entered in every hit or log line in the log file that indicates the status of a file/page request.

Information codes 1XX - This code indicates a provisional response, consisting only of the Status-Line and optional headers, and is terminated by an empty line. It is usually only for experimental purposes.
100 Continue
101 Switching Protocols

Successful 2XX - This status code indicates that the client's request was successfully received, understood, and accepted. Possible Successful Status Codes are:
200 OK
201 Created
202 Accepted
203 Non-Authoritative Information
204 No Content
205 Reset Content
206 Partial Content

Redirection 3xx - This code indicates that further action needs to be taken by the user agent in order to complete the request.
300 Multiple Choices
301 Moved Permanently
302 Moved Temporarily
303 See Other
304 Not Modified
305 Use Proxy
306 Unused
307 Temporary Redirect

Client Error 4XX - Errors that occurred due to an invalid or failed request by the client or visitor's browser. Possible client error codes are:
400 Bad Request
401 Unauthorized
402 Payment Required
403 Forbidden
404 Not Found. This is the most common error on the reports, and means that someone has requested an object that is not within your site. This could be because of:
Bad links to pages within your site.
Links within a page to a graphic that does not exist.
Someone misspelling the URL to get to an object on your site.
Outside links to your site that is out of date...
405 Method Not Allowed

406 None Acceptable
407 Proxy Authentication Required
408 Request Timeout
409 Conflict
410 Gone
411 Length Required
412 Precondition Failed
413 Request Entity Too Large
414 Request-URI Too Long
415 Unsupported Media Type
416 Requested Range Not possible
417 Expectation Failed

Server Error 5XX - Errors that occurred due to an invalid or failed task on the server. Possible server error codes are:
500 Internal Server Error
501 Not Implemented
502 Bad Gateway
503 Service Unavailable
504 Gateway Timeout
505 HTTP Version not supported

Internal Errors 6XX - Errors that occurred in the link checking program itself (these are unofficial error codes used to report problems).
600 Malformed URI
601 Connection Timed Out
602 Unknown Error
603 Could Not Parse Reply
604 Protocol Not Supported

SSL or Secure Sockets Layer - A proposed open standard developed by Netscape Communications for establishing a secure communications channel to prevent the interception of critical information, such as credit card numbers. The primary purpose of SSL is to enable secure electronic financial transactions on the World Wide Web, although it is designed to work with other Internet services as well. This technology, which uses public key encryption, is incorporated into many Web browser and Web servers. Refer to http://www.rsa.com for more on public key encryption and the subject of cryptography.

Server - Also called host. This is a computer that offers services on a network. On the World Wide Web, a server is a computer that runs the Web server software that responds to HTTP protocol requests.

Server Error Codes - Status Error Codes in the 5XX range indicating the type of errors that occurred during a file/page request. See Status Codes

Shared Client Session - A Session, which is, shared among 2 or more views i.e. a client who has visited more than 1 Page. Other Log File Analyzers shows this as separate clients, which do not tally with the total clients. Some show it as "Not Applicable"

Spiders or robots - An automated program, which searches the Internet and add data to search engines. On a web site the robots.txt file specifies which directories the spider should catalog.

Time Zone - This is the time relative to Greenwich Mean Time (0:00), e.g. GMT - 6 for Central America. Most World Wide Web servers record hits or requests in local time where the server is located.

URL - Uniform Resource Locator is a means of identifying an exact location on the Internet. An URL consists of four parts:
Protocol Type (HTTP),
Machine Name (domain),
Directory Path (/), and
File Name (default.htm),
e.g. http://www.surfstats.com/surfstats/default.htm

User (client or visitor) - Anyone who visits the site at least once. A user is uniquely identified by an IP address. A unique visit is the number of unique visits in a specific time period from a unique IP Address.

User (client) Session - A series of consecutive requests from a user to an Internet site. A user session is terminated when a user does not make another request for more than 30 minutes.
If for example a visitor with IP Address 1.2.3.4, visits the site, logs out and another visitor logs in an hour later with the same IP Address, there would be two user sessions but one unique visitor.

Unique visitors - Unique visitors are counted using the visitor's IP address or cookie information to identify the visitor.

Country Internet Codes

Country Internet Codes

Top Level Domains or ISO Internet Codes are the suffix added to a domain name, e.g. myname.co.**us**. Following is a list of country codes.

.aero Aviation .biz Business Organizations .com Commercial .coop Co-Operative Organizations .edu Educational .gov US Government .info Open TLD .int International Organizations .mil US Dept of Defense .museum Museums .name Personal .net Networks .org Organizations .ac Ascension Island .ad Andorra .ae United Arab Emirates .af Afghanistan .ag Antigua and Barbuda .ai Anguilla .al Albania .am Armenia .an Netherlands Antilles .ao Angola .aq Antarctica .ar Argentina .as American Samoa .at Austria .au Australia .aw Aruba .az Azerbaijan .ba Bosnia and Herzegovina .bb Barbados .bd Bangladesh .be Belgium	.bf Burkina Faso .bg Bulgaria .bh Bahrain .bi Burundi .bj Benin .bm Bermuda .bn Brunei Darussalam .bo Bolivia .br Brazil .bs Bahamas .bt Bhutan .bv Bouvet Island .bw Botswana .by Belarus .bz Belize .ca Canada .cc Cocos (Keeling) Islands .cd Congo, Democratic republic of the (former Zair .cf Central African Republic .cg Congo, Republic of .ch Switzerland .ci Côte d'Ivoire .ck Cook Islands .cl Chile .cm Cameroon .cn China .co Colombia .cr Costa Rica .cu Cuba .cv Cape Verde .cx Christmas Island .cy Cyprus .cz Czech Republic (former Czechoslovakia)

.de Germany .dj Djibouti .dk Denmark .dm Dominica .do Dominican Republic .dz Algeria .ec Ecuador .ee Estonia	.hn Honduras .hr Croatia .ht Haiti .hu Hungary .id Indonesia .ie Ireland .il Israel .im Isle of Man

.eg Egypt	.in India
.eh Western Sahara	.io British Indian Ocean Territory
.er Eritrea	.iq Iraq
.es Spain	.ir Iran
.et Ethiopia	.is Iceland
.fi Finland	.it Italy
.fj Fiji	.je Jersey
.fk Falkland Islands	.jm Jamaica
.fm Micronesia	.jo Jordan
.fo Faroe Islands	.jp Japan
.fr France	.ke Kenya
.ga Gabon	.kg Kyrgyzstan
.gb United Kingdom	.kh Cambodia
.gd Grenada	.ki Kiribati
.ge Georgia	.km Comoros
.gf French Guiana	.kn Saint Kitts and Nevis
.gg Guernsey	.kp Korea, Democratic People's Republic of
.gh Ghana	.kr Korea, Republic of
.gi Gibraltar	.kw Kuwait
.gl Greenland	.ky Cayman Islands
.gm Gambia	.kz Kazakhstan
.gn Guinea	.la Lao People's Democratic Republic
.gp Guadeloupe	.lb Lebanon
.gq Equatorial Guinea	.lc Saint Lucia
.gr Greece	.li Liechtenstein
.gs South Georgia and the South Sandwich Islands	.lk Sri Lanka
.gt Guatemala	.lr Liberia
.gu Guam	.ls Lesotho
.gw Guinea-Bissau	.lt Lithuania
.gy Guyana	.lu Luxembourg
.hk Hong Kong	.lv Latvia
.hm Heard and McDonald Islands	.ly Libyan Arab Jamahiriya

.ma Morocco	.pf French Polynesia
.mc Monaco	.pg Papua New Guinea
.md Moldova	.ph Philippines
.mg Madagascar	.pk Pakistan
.mh Marshall Islands	.pl Poland
.mk Macedonia	.pm St. Pierre and Miquelon
.ml Mali	.pn Pitcairn
.mm Myanmar	.pr Puerto Rico
.mn Mongolia	.ps Palestine.pt Portugal
.mo Macau	.pw Palau
.mp Northern Mariana Islands	.py Paraguay
.mq Martinique	.qa Qatar
.mr Mauritania	.re Reunion
.ms Montserrat	.ro Romania
.mt Malta	.ru Russia
.mu Mauritius	.rw Rwanda
.mv Maldives	.sa Saudi Arabia
.mw Malawi	.sb Solomon Islands
.mx Mexico	.sc Seychelles
.my Malaysia	.sd Sudan
.mz Mozambique	.se Sweden
.na Namibia	.sg Singapore
.nc New Caledonia	.sh St. Helena
.ne Niger	.si Slovenia
.nf Norfolk Island	.sj Svalbard and Jan Mayen Islands
.ng Nigeria	.sk Slovakia
.ni Nicaragua	.sl Sierra Leone
.nl The Netherlands	.sm San Marino
.no Norway	.sn Senegal

.np Nepal .nr Nauru .nu Niue .nz New Zealand .om Oman .pa Panama .pe Peru	.so Somalia .sr Surinam .st Sao Tome and Principe .su USSR (former) .sv El Salvador .sy Syrian Arab Republic .sz Swaziland

.tc The Turks & Caicos Islands .td Chad .tf French Southern Territories .tg Togo .th Thailand .tj Tajikistan .tk Tokelau .tm Turkmenistan .tn Tunisia .to Tonga .tp East Timor .tr Turkey .tt Trinidad and Tobago .tv Tuvalu .tw Taiwan .tz Tanzania .ua Ukraine .ug Uganda .uk United Kingdom .um United States Minor Outlying Islands .us United States	.uy Uruguay .uz Uzbekistan .va Holy See (Vatican City State) .vc Saint Vincent and the Grenadines .ve Venezuela .vg Virgin Islands British .vi Virgin Islands U.S .vn Vietnam .vu Vanuatu .wf Wallis and Futuna Islands .ws Samoa .ye Yemen .yt Mayotte .yu Yugoslavia .za South Africa .zm Zambia .zw Zimbabwe