# Roxen™ User Manual

roxen
Campaign

LogView

This chapter contains information on different ways of using LogView to view the extracted statistics.

Using LogView

LogView lets the user view the gathered statistics in an ordinary web browser. If LogView has been installed on the Roxen virtual server with the URL

    http://hostname/intranet/

the default URL to the LogView configurations page is

    http://hostname/intranet/logview/

However, the administrator may choose another location within the virtual file system. The configurations page can also be found under the *Plugins/LogView* tab in SiteBuilder.
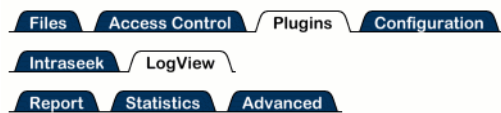
The Plugins/LogView tab

If the administrator has specified that access to the LogView pages should be password protected, the user will get a pop-up dialog where he has to log in before he can access the LogView page.

On the top of all the LogView pages is located a button menu which helps the user to navigate between the four pages *Statistics, Report and Advanced*. These pages provide different ways of looking at the gathered statistics.

The main difference between the *Statistics, Reports* and *Advanced* pages lies in what possibilities the user has to affect the presentation. The *Reports* page contains a number of reports that has been prepared by the administrator, and is therefore the most static one - all the user needs to do is to select which report he wishes to see. The *Statistics* page is the easy way to create reports, and provides basic possibilities to control what statistics to see and in which form, and in the *Advanced* page the user has even more freedom to control the details.

These three levels of control possibilities makes it possible to provide quick-and-easy access for the users who either does not need full control or can have the administrator or site creator tailor a report for them, while the users that really need full control over the details can still get it.

Logging sites

When someone retrieves a page from a web site, all the actions this person does is saved to a file in a special log format. Every action from downloading files, fetching HTML-pages with pictures to what web-browser or operating system the user has is saved as plain text in this file. This file can later be analyzed to find out various behaviors the site's users have.

The process of building a log file works something like this: the users browser sends requests to your Web server for various resources. These resources may include HTML, graphics, audio, and whatever other types of files are on the site. Upon receiving the request from a browser, your Web server accesses the file and sends it back to the client browser.

A lot of useful information is buried in this file. Information about where bottlenecks and errors occur in the server or on the site or where the most popular features on the site are.

The information gathered in the log files can as all statistics information be made to fool those not as attentive or those who don't have any knowledge about the area the statistics is for. Here is a screen-shot that there is obviously something fishy about.
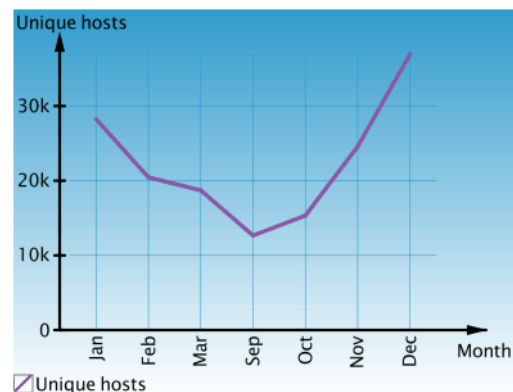
It is easy too fool those not paying attention

To make the best use of statistics information it is important to know what all functions and special terms concerning LogView means. Hence, a compilation of all the important terms and functions in LogView can be found in the *Terms and functions* page.

Statistics

The statistics page provides an easily overviewed and almost self-explanatory way of looking at statistics gathered by all the functions that are enabled on the current statistics group.

Selecting statistics group

In the field labeled *Group* the current statistics group is shown. If several groups are defined, the user can select one of them in a list box. By default only one group is defined. It takes its data from the log of the Roxen virtual server that LogView is installed on, and the name of this group is the same as the name of the server.

If no hits have been received since the creating of the selected statistics group, this is shown by the message "No items are logged". This should be remedied within one hour after the site has been accessed.

Select group and period of time

Selecting period of time

In the *Period of time* box, the period of time for which to show statistics can be specified. The currently selected period of time is displayed within parentheses right next to the label on the top of the box. The highlighted periods inside the box are the ones for which logged data is available, and clicking one of them simply selects that period of time. By default only years and months are shown, but once a month is selected the box is expanded to also contain weeks and days.

On the *Statistics* page, statistics can only be shown for a single year, month, week or day. To specify a more exact time interval, the user has to go to the *Advanced* page.

Selecting kind of statistics

Below the *Period of time* box, a number of available functions are listed in groups. To see statistics of a certain kind the user just has to select the corresponding function in the list. An overview over the functions can be found in the *Terms and functions* page.

Select statistics function

Selecting format

Once the user has selected the kind of statistics to show, a report is displayed at the bottom of the page. This report is in a default format that is dependent of the kind of statistics that was selected. If the user is not pleased with this format, he/she can choose another one in the list box in the title of the report. Below is a selection of statistics reports for www.roxen.com under week 47, 1998. These reports selected by first choosing the www.roxen.com group, then clicking on 1998, Nov and last week 47.
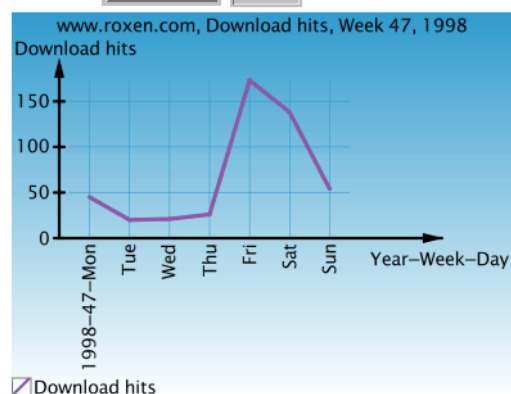
Download hits - line chart

This report is selected by clicking on the function *Download hits* and then choosing *line chart* from the pull down menu.

Download hits - bar chart

This report is selected by clicking on the function *Download hits* and then choosing *bar chart* from the pull down menu.
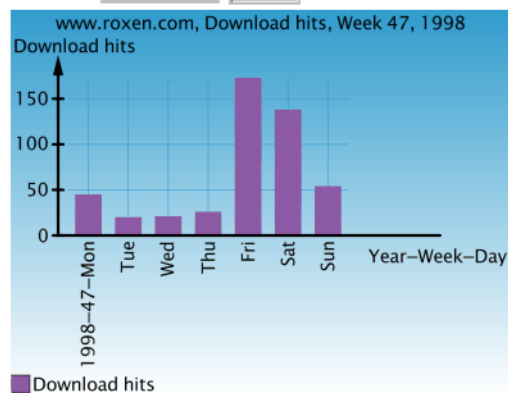

External links, pages - table

This report is selected by clicking on the function *External links, pages* and then choosing *table* from the pull down menu.
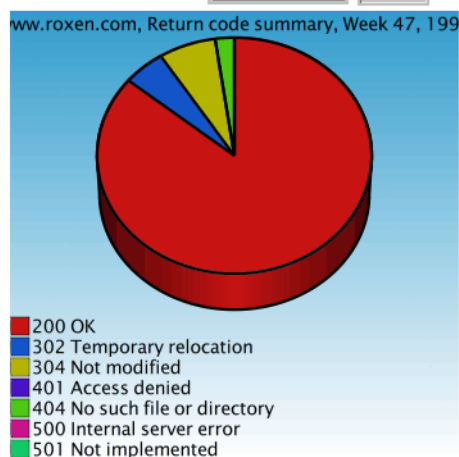
Return code summary - 3D pie chart

This report is selected by clicking on the function *Return code summary* and then choosing *3d pie chart* from the pull down menu.

Reports

The report page is used for viewing reports that has been prepared by the LogView administrator.

Advanced

In this page, the user can control all of LogView's capabilities for viewing statistics.

The advanced statistics page

Selecting group

Select statistics group

As was the case on the *Statistics* page, the statistics group, if more than one is available, can be selected in a list box.

Selecting report

Select type of report

The user can select the kind of report from all the kinds that are available for the selected statistics group in the first list box on the row.

The list box after *as* default labeled *Day* is used to specify the granularity of the diagram or table. If *days* is selected here, a resulting bar chart will have one bar for each day for

which data is available, and a table will accordingly have one row for each day. If instead *month* is selected, the bar chart will have one bar for each month, with a height that is the the sum of the all the values of the days of this month.

The list box labeled *per* selects what should be printed for example as labels on the bars of a bar chart. To get a chart where each day has it's own bar and the label Year-Week-Day (e g 1998-25-Tue) the user can select the report to be shown "as day, per week". A report that is "as day, per month", has the label Year-Month-Day (e g 1998-Jun-16) on its bars instead.

Trying to set the *as* value larger than the *per* value gives the following error message:

```
LogView Error: unit cannot be greater than
per
```

Setting both values equal has an interesting effect when using the Sum feature, as we will see in the next section.

Selecting presentation

Select presentation

In the first list box, the user can choose between *Append* or *Sum*. *Append* is the default, and gives sums for each day, week, month or year. Using the *Sum* feature changes this behavior somewhat - instead of summing the data inside each day, month or year the data for several such units can be summed, to answer questions like "how many hits were processed in total between the 1998-06-02 and 1998-07-11?" However, to use this feature we have to use the trick to set the *as* and *per* time periods the same.

The two pictures below shows the difference in *Append* and *Sum* when the setting is month per month. The report is *Search engines* and the time period is between December, 1998 and January, 1999.
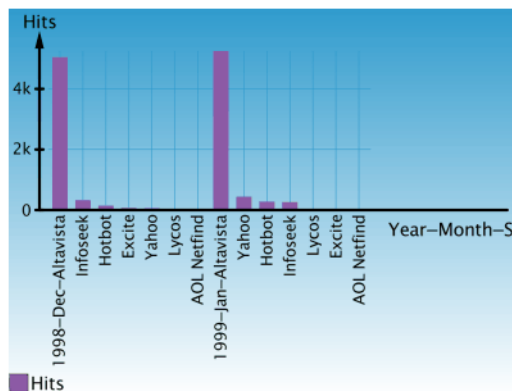
Append
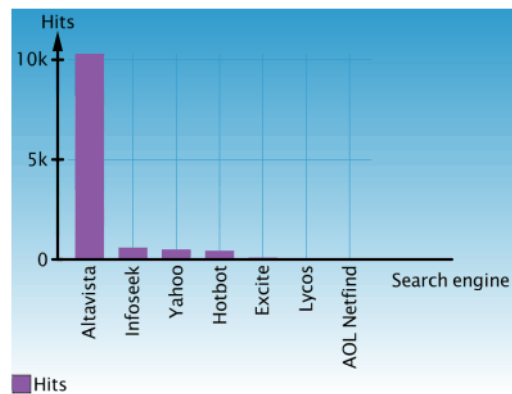
Sum

Another scenario is when the user has set the report to days per week, and the time period is three weeks, where *Append* will present it as three weeks in a row. Sum will instead present it as one week, where each weekday is the sum of that weekday in all three weeks. Due to lack of space no screen-shots of this scenario is possible in the manual.

There is also a possibility to set the maximum number of rows to be shown. The term rows seem to imply the use of a table, but the limitation works on all kinds of diagrams as well, and then limits the number of bars, pie slices or other corresponding items.

Selecting from and to values

Select time period(TU)

The user can specify the exact time period for which statistics should be shown by selecting the years, months and days for the beginning and end of the period.

Selecting display

Select display type

When the user has specified all of the above, he can start the display process by pressing one of the display buttons in the row at the bottom of the page. These are divided into categories by their form of presentation.

Tables:

o HTML table

#	Year	Month	Hits
1	1998	Sep	469 396
2	1998	Oct	548 271
3	1998	Nov	787 083
4	1998	Dec	933 046
5	1999	Jan	693 880
6	1999	Feb	738 378
7	1999	Mar	770 851
8	1999	Apr	676 930
9	1999	May	632 297
10	1999	Jun	785 343
11	1999	Jul	610 257

Hits at www.roxen.com, September 1998 through July 1999

Settings: Hits as Month per Year, Append

o ASCII table

```
#      Domain          Pages      Hits
----------------------------------------
1      unknown         196459     1271793
2      aol.com         20943      206164
3      liu.se          28682      194564
4      ne.jp           11034      96744
5      telia.com       10273      92487
6      home.com        11431      91798
7      swipnet.se      9303       89153
8      t-online.de     10559      88394
9      uu.net          13339      88301
10     ac.at           40203      82999
11     co.uk           9681       78069
12     algonet.se      17114      72942
13     mxs.nl          5718       72617
14     co.jp           7658       51506
15     prognet.com     8037       51372
16     or.jp           7186       44082
17     com.au          9074       36041
18     com.br          3423       35499
19     ac.uk           3619       33242
20     rr.com          5217       33204
```
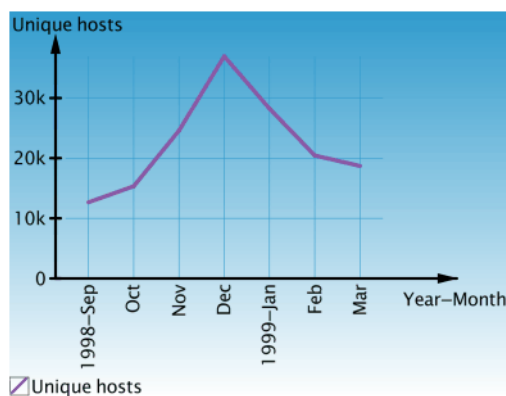
Most active domains at www.roxen.com, September 1998 through July 1999

Settings: Most active domains as Month per Year, Sum

Charts:

o Line chart

Unique hosts at www.roxen.com, September 1998 through March 1999

Setting: Unique hosts as Month per Year, Append

o Bar chart
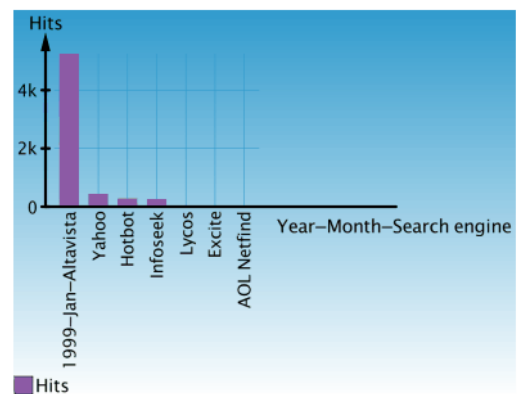
Search engines at www.roxen.com, January 1999

Settings: Search engines as Year per Month, Append

o Sum bars chart
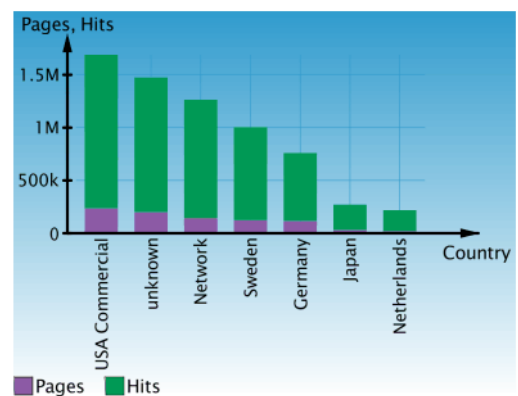
Most active countries at www.roxen.com, September 1998 through July 1999

Settings: Most active countries as Month per Month, Sum, max 7 rows

o 2D pie chart
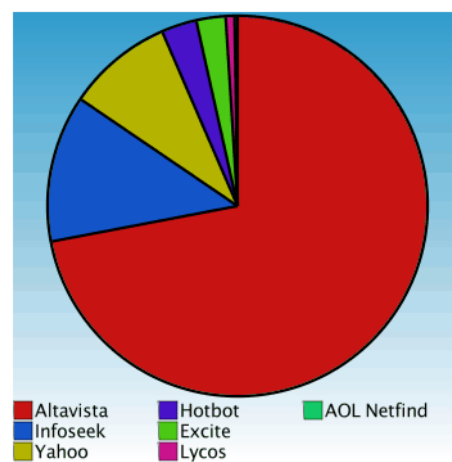
Search engines at www.roxen.com, January 1999 through July

1999

Settings: Search engines as Month per Month, Sum

o 3D pie chart

Hits at www.roxen.com, September 1998 through July 1999

Settings: Hits as Month per Month, Append.

Miscellaneous:

o Map

A world map, nice for displaying statistics over Most active *countries*. Not available in any other case. Below is a slightly cut version of the map, made for this manual.

Most active countries at www.roxen.com, September 1998 through July 1999

Settings: Most active countries as Month per Year, Append

o Export

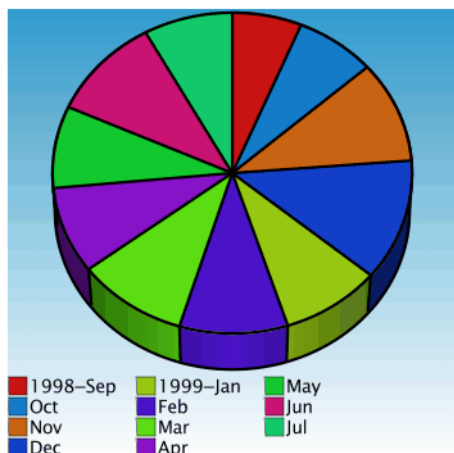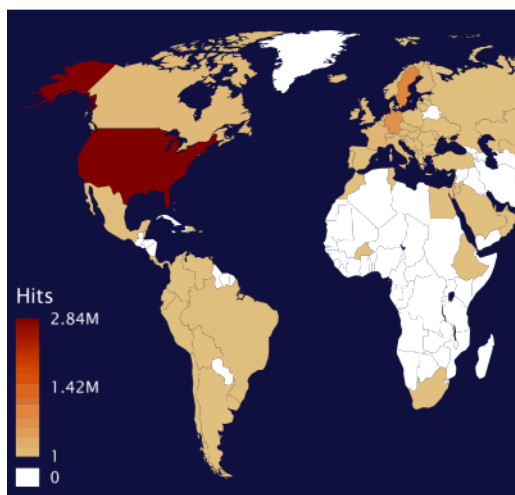Exports the table in a tab/newline separated format, suitable for MS Excel.

o Tag

Displays the LogView RXML tag with all the parameters (see appendix B) that generates the created report. Great for cut-and-paste!

```
<logview  group='www.roxen.com'
  report='Most active countries'
  unit=month  per=year
  op=sum  max=100
  from-year=1998
  from-month=9
  from-day=1
  to-year=1999
  to-month=7
  to-day=31
  display=world-map>
```

Most active countries at www.roxen.com, September 1998 through July 1999

Settings: UTSL! (Use the source Luke!) or read the tag-code and find out the right settings.

Terms and functions

Terms

This part describes the difficult words and terms in Log-View. These words are either web analysis terms or used by LogView to describe it's functions.

Bandwidth In a general sense, this term describes information-carrying capacity. In LogView this applies to how much data (in bits per second) that has been transferred from the server.

Hit A hit is a request from a browser to the server. Every element of a requested page, text, pictures, and interactive items is counted as one hit to the server. A normal web page contains an average of at least six elements, i.e. six hits. Hits is not the preferred way of measuring traffic on a site as the number of hits often varies widely.

Page load Page load aka page views or page deliveries gives the number of times a web page has been requested. Page load, not hits, are the preferred counting method site-traffic estimates and measurement.

Search engine The search functions can identify the following search engines: AOL Netfind, Alta Vista, Excite, Infoseek, Hotbot, Lycos and Yahoo.

Session A session is tied to a hostname where the hits has occured up to 20 minutes between each other.

Period of time/time period In LogView reports can be created to show the data between two points of time. What is in between those two points of time is the period of time.

User variable Variables which the administrator or the privileged user can set in one of the LogView wizards found in the configurations interface. Some of the functions makes use of these variables when monitoring the log files. This term is only used when describing some of the functions below.

Functions

This part describes all the functions used to present data about the site in LogView.

Site activity

This group of functions monitors the overall activity on the site.

Hits Gives a summary of the number of hits during the period of time.

Page load Summaries the number of hits per time period and page.

Bandwidth Calculates the average number of sent bits per second and time period.

Visitor origins

This group of functions monitors where hits originates from.

Most active countries This function describes which countries are generating the most hits.

Most active domains This function tells which internet domains are responsible for generating most hits. A domain can be roxen.com, foobar.net, etc.

Most active hosts This term gives a summary of each hosts hits. The most active host is the one who has the most hits. A host can e.g., be an IP-number.

Authentications

Gives information about user logins.

Authenticated users Gives a summary of the number of hits/time period and authenticated user.

Authenticated hosts & users Gives a summary of the number of hits/time period, hostname and the authenticated user.

Popular contents

This group of function monitors what has been requested on the site.

Popular pages Gives a summary of hits/time period and page.

Popular non-pages Gives a summary of hits/time period and non-page, where non-pages are all files which are not pages e.g., a zip-file or a pdf-document.

Popular directories Gives a summary of hits/time period and directory.

Visitor sessions

This group of functions gives information about visitors coherent sessions.

Average hits & pages per session This function calculates the average number of hits and the average number of pages per time period and session.

Average session length This functions calculates the average session length in minutes per TU.

Most frequent entry pages Summaries the number of hits per time period and page where the page is the first page accessed in a session.

Most frequent exit pages Summaries the number of hits per time period and page where the page is the last page accessed in a session.

Visitor profiles

This group of functions gives miscellaneous information about the users.

Most common OS Summaries the number of hits per operating system and time period.

Most common browser Summaries the number of hits per browser and time period.

Most common browser & version Summaries the number of hits per unique browser version and time period.

Most common browser, version & OS Summaries the number of hits per unique browser version, it's operating system and time period.

Pathways

This group of functions monitors links that points towards the site.

Internal links, pages Summaries the number of hits to the linked page (the "to" page) where the link-page (the "from" page) was an internal page. Pages that match the user variables *Server Domain Name* or *Server Alias* are counted as internal pages e.g., all pages on *your.site.com* are internal pages.

External links, pages Summaries the number of hits to the linked page where the link page was an external page.

External links, non-pages Calculates the number of hits to files which are not pages e.g., zip-files, pdf documents, etc.

Search

This group of functions gives information about visitors who hahave found the site using search engines.

Search hits Summaries the number of hits where the previous page was a search engine.

Search words Summaries the number of words in the search string per time period where the previous page was a

search engine. The stopwords 'and', 'or', 'not', 'the', 'an', 'a', 'is' and 'are' are not counted.

Search phrases Summaries the number of phrases per time period where the previous page was a search engine and the phrase a search string.

Search engines Summaries the number of hits per search engine where the previous page was a search engine.

Download

This group of functions monitors which downloads have been made.

Download hits Summaries the number of hits per time period where the filename matches the uservaiable *Download files* i.e., total number of downloads.

Popular files Summaries the number of hits per time period and filename, where the filename matches the user variable *Download files* i.e., sorted by files.

Maintenance

This group of functions gives information necassary for troubleshooting.

Return code summary Summaries the number of returned status codes.

Error code specification Summaries the number of returned status codes per time period and filename, where the status code is an error code i.e., separated from 200 and the interval 300-399.

Extensions

This function monitors the number of unique hosts who have accessed the site.

Unique hosts Summaries the number of hits per time period and unique host, where a host can be represented by a hostname or an IP-number. It should be noted that both hostname and IP-number can represent the same host, which implies that they will be counted as two different hosts. This is a known bug.

IntraSeek

How to search with IntraSeek

This text can be added to a web page to tell the users how to use the IntraSeek search form.

It is possible to use quotation marks (?) and asterisks (*) to broaden searches. A search for "net*" might match "netscape", "nethack", "network" and so on. A search for "int??net" matches "intranet" as well as "internet". Note that IntraSeek requires that the user specifies at least three characters in front of the "*" notation, and that there is no difference between lower- and uppercase searches.

You can use quotation marks to search for a phrase. For example, a search for "John Carl Smith" will search for persons with this name. Without quotes, you would have get any pages that use any of those common names.

Boolean

Boolean search allows you to include or exclude documents containing certain words through the use of the operators AND, OR, NOT and XOR.

If you just type some keywords, the "OR" operator is assumed. For example, searching for "coca-cola pepsi jolt" will accept all documents containing one or more of any of these three words.

Searching for "coca-cola AND pepsi" only displays the documents that contain BOTH coca-cola and pepsi.

Searching for "(coca-cola AND pepsi) AND NOT jolt" will display all documents that contain the words "coca-cola" and "pepsi", but will leave out documents containing anything about "jolt".

You can add any level of "(" and ")" to build evaluation trees.

Searching for "(coca-cola XOR pepsi)" will display all documents that contain the word "coca-cola" or the word "pepsi", but not both.