

## Web metrics with Webalizer

Webalizer is included in most Linux distributions. If not in your release, you can find this at:  
<http://webalizer.linux-mirror.org/>

For this document, we'll be using the version available on RHEL 5: webalizer-2.01\_10-30.1

To get this started, install the RPM downloaded from the above site:  
Rpm -ivf webalizer-2.01\_10-30.1.rpm

This will install the binaries and a default configuration file at /etc/webalizer.conf. The webalizer statistics may contain sensitive information. To prevent unauthorized people from viewing, webalizer default service is on localhost:

This configuration file maps the webalizer log analysis results (generated daily) into the URL space. By default these results are only accessible from the local host.

```
Alias /usage /var/www/usage
```

```
<Location /usage>  
  Order deny,allow  
  Deny from all  
  Allow from 127.0.0.1  
  Allow from ::1  
  Allow from .example.com  
</Location>
```

Many times, a more useful setup is to allow usage access remotely if over https and with an authorized user. Here's a configuration to do this:

```
Alias /usage /var/www/usage/  
<Directory "/var/www/usage/">  
  Options None  
  Order allow,deny  
  Allow from all  
  AuthName "Usage Access"  
  AuthType Basic  
  AuthUserFile /var/www/etc/htpasswd.users  
  Require valid-user  
</Directory>
```

Once the configurations are added to Apache, restart the service:  
➤ service httpd restart

Webalizer is run through the daily cron job: /etc/cron.daily/00webalizer and looks for an access log in /var/log/httpd/access.log. If you have a different log name convention, you will need to modify this job to look at the correct log.

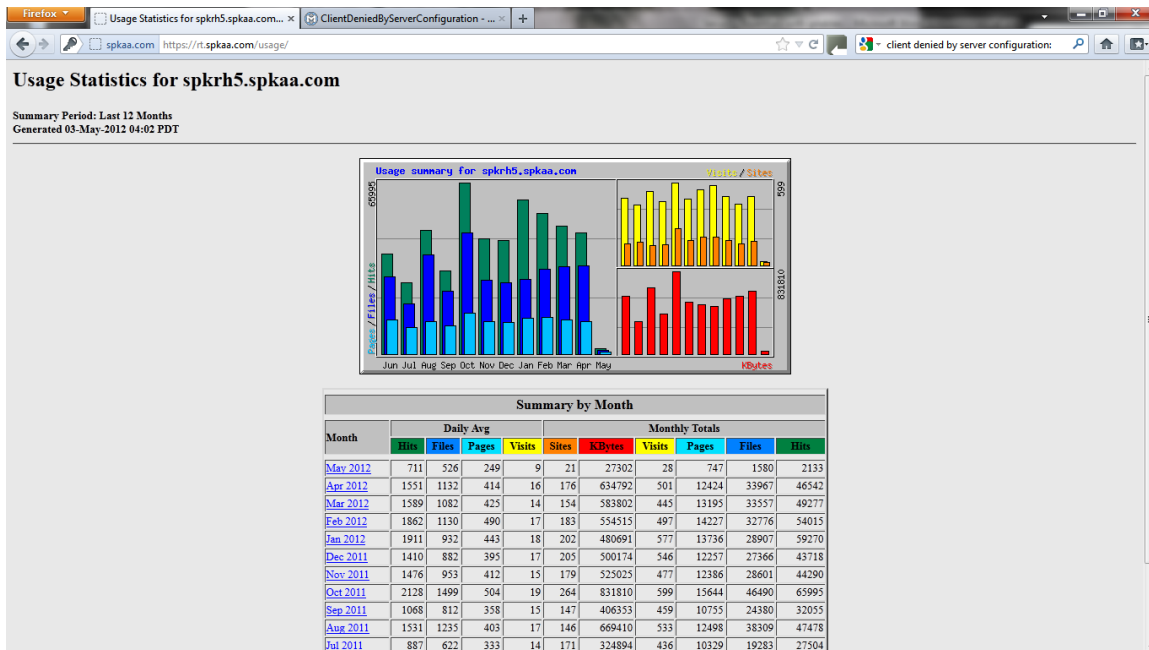
For larger scale environments, you likely have multiple hosts all logging to separate logs. To allow for these to be merged into one log for stats, we use the logrrsolvemerge.pl utility from AW Stats:

[http://awstats.sourceforge.net/docs/awstats\\_tools.html](http://awstats.sourceforge.net/docs/awstats_tools.html)

Usage is simple:

```
./logrrsolvemerge.pl host1.log host2.log host3.log > access.log
```

Here's a few screen prints on one of our non-production environments:





www.spkaa.com  
Ph: 888-310-4540

SPK and Associates  
900 E Hamilton Ave, Ste. 100  
Campbell, CA 95008

Here's the next screen if you select May 2012:

Usage Statistics for spkrh5.spkaa.com

Summary Period: May 2012  
Generated 03-May-2012 04:02 PDT

[\[Daily Statistics\]](#) [\[Hourly Statistics\]](#) [\[URLs\]](#) [\[Entry\]](#) [\[Exit\]](#) [\[Site\]](#) [\[Referrers\]](#) [\[Search\]](#) [\[Users\]](#) [\[Agents\]](#) [\[Countries\]](#)

Monthly Statistics for May 2012		
Total Hits		2133
Total Files		1580
Total Pages		747
Total Visits		28
Total KBytes		27302
Total Unique Sites		21
Total Unique URLs		292
Total Unique Referrers		53
Total Unique Usernames		1
Total Unique User Agents		36
	Avg	Max
Hits per Hour	29	229
Hits per Day	711	1130
Files per Day	526	864
Pages per Day	249	400
Visits per Day	9	18
KBytes per Day	9101	15328
Hits by Response Code		
Code 200 - OK		1580
Code 301 - Moved Permanently		335
Code 302 - Found		75
Code 304 - Not Modified		140
Code 400 - Bad Request		1