
How to set up a Web Counter using Python

by

Prof. Michael Morrell
mgm@nmsu.edu

New Mexico State University.
Department of Engineering Technology.

This TechTip presents an easy way to set up counters for each web page in your web directory on a unix or linux system. Instead of having your anchors point to html pages, those anchors can point to one python script which can be used to handle all the pages. Set-up your anchors to point to the cgi script that will count the number of times that page is accessed. The cgi counter script can record the access event in a log book with the date, time and if desired, the IP address of the requestor.

A very simple script to do this is shown below.

```
LISTING FOR index.cgi
-----
#!/usr/bin/python
import cgi,os,time
env=os.environ
try:
    ip=env['REMOTE_ADDR']
except:
    ip='localhost'
form=cgi.FieldStorage()
try:
    f=form['f'].value
except:
    f='index'
ti=time.localtime()
t = "%04d/%02d/%02d %02d:%02d %s %s" % (ti[0],ti[1],ti[2],ti[3],ti[4],ip,f)
log=open("log","a")
log.write("%s\n" % t)
dd = open(f).read()
print "Content-type: text/html\n\n"
print dd
# End of Python Script
```

Please Note: If you have not used python before, the indentations shown after a line ending with a colon (:) are 4 space indents. (Don't Use TABS). If you indent two levels, use eight spaces, etc.

To set this up, make sure that the first line of the program is the `#!/usr/bin/python` line. If your python executable is not located there, then you must figure out where it is located, and put that path as the first line, preceded by the `#!` code. This tells the operating system and python to run that program when this file is requested.

We recommend you name this file `index.cgi`. The cgi script must be installed in the same directory as the html files that are being counted.

Let's say you have the following files in your directory.

page1, page2, index, page3, and page4.

Anchors for the above would be like as follows

```
<a href=index.cgi?f=page1 > Page 1 </a>
<a href=index.cgi?f=page2 > Page 2 </a>
<a href=index.cgi?f=page3 > Page 3 </a>
<a href=index.cgi?f=page4 > Page 4 </a>
<a href=index.cgi?f=index > Index </a>
```

This would produce log entries like:

```
2007/01/05 08:20 128.134.23.5 page1
2007/01/05 08:23 128.134.23.5 page2
2007/01/05 08:24 128.134.23.5 page3
2007/01/05 08:24 128.134.23.5 page4
2007/01/05 08:28 128.134.23.5 index
```

To enable scripting on your webserver, requires you modify your webserver configuration file often called /etc/httpd/httpd.conf or /etc/httpd/conf/httpd.conf to have the setting ExecCGI as a permission in the directory you will store this script into.

Also, you must set the executable bit in this index.cgi script. (use chmod command)
AND you must set the write bit on the log. (using chmod)

You can do these as follows.

```
chmod a+x index.cgi
touch log
chmod a+w log
```

You may need to ask your system administrator to help with this.

You can easily read the results of the log file, with summaries with another script called logRead.cgi. This script can be located in the same directory as the html files, and the python script we just described. When you type in a URL to this python script, it will show you a page with a neat listing of the log file created by the index.cgi python script describe previously. An example of the URL to run this python script would be:

<http://www.yourdomain.edu/myfiles/logRead.cgi>

LISTING FOR logRead.cgi

```
-----  
#!/usr/bin/python  
import cgi  
l=open("log").read().split("\n")  
pages = {}  
for i in range (len(l)):  
    if (l[i] != ""):  
        x = l[i].split()  
        try:  
            pages[x[3]]=1 + pages[x[3]]  
        except:  
            pages[x[3]] = 1  
k=pages.keys()  
k.sort()  
print "Content-type: text/html\n\n"  
print "<html><body><table border=1>"  
for i in range (len(k)):  
print "<tr><td>%s</td><td align=right>%d</td></tr>" % (k[i],pages[k[i]])  
print "</table></body></html>"  
# End of Python Script  
-----
```

For more reading please refer to the following.

Python: <http://www.python.org>

Apache Web Server: <http://httpd.apache.org/docs/1.3/misc/FAQ.html>

CGI Scripting: <http://www.ietf.org/rfc/rfc3875>

Good Python Book: Lutz, Mark. "Programming Python". 2001. Second Edition. O'Reilly Publishing. California.