



# **SANS Institute**

## Information Security Reading Room

# **Using Open Source Software to Proxy, Authenticate, and Monitor User Web Habits**

---

Jason Gregg

Copyright SANS Institute 2019. Author Retains Full Rights.

This paper is from the SANS Institute Reading Room site. Reposting is not permitted without express written permission.

**Jason D. Gregg**

**GSEC Practical Assignment**

**Version 1.2f**

**Using Open Source Software to Proxy, Authenticate, and Monitor  
User Web Habits**

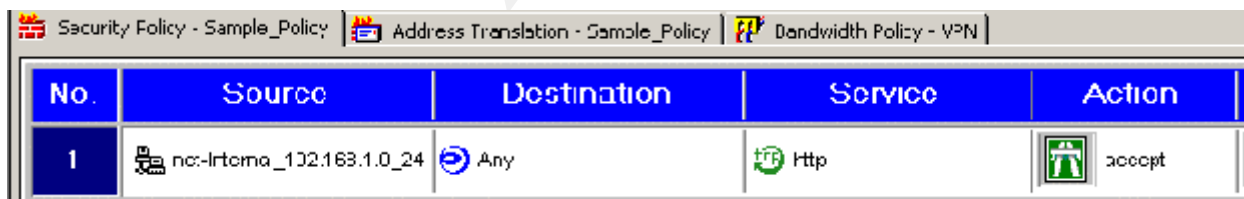
© SANS Institute. All rights reserved. Author retains full rights.

## Introduction

In today's corporate environment employee access to Internet resources is, for many, a daily job function. Fundamentally, the web is utilized for simple day-to-day operations including virus updates, software patches and hardware drivers and other network administrator duties. Other divisions of an organization may require the World Wide Web for resources such as market updates, company news, and general research on products or competitor's products. In today's business environment, access to the World Wide Web is no longer an option for many companies, but instead a requirement. Such global World Wide Web access offered to employees raises issues and concerns over user web habits and issues with users utilizing the World Wide Web for business use during business hours. This paper will attempt to address what time and again is a problem for network and security administrators: monitoring user access to the Internet in an environment where blocking resources may not be ideal, cost effective, or in accordance with company policy.

## The Issue with Stateful Inspection Firewalls

A large number of organizations today have replaced the more traditional proxy firewall design for the performance and flexibility a stateful inspection design firewall<sup>1</sup>. Such a change has fundamentally altered the network services organizations are willing to allow into their corporate environment and have, in many cases, limited administrators abilities to quickly and easily access information related to user web habits. In a stateful inspection firewall such as Checkpoint Firewall-1<sup>2</sup>, a standard web browsing rule might look like the following:



The screenshot shows a configuration window for a Checkpoint Firewall rule. The window title is "Security Policy - Sample\_Policy". The rule is named "1" and is configured as follows:

No.	Source	Destination	Service	Action
1	no-Intoma_132.163.1.0_24	Any	Http	accept

This rule would be logged just like any other rule in the ruleset with the source and destination IP address, service, protocol and various other relevant log data. The following screenshot illustrates Checkpoint's log entries:

Type	Action	Service	Source	Destination	Proto.
log	accept	www-http	192.168.1.60	209.133.65.45	tcp
log	accept	www-http	192.168.1.60	207.188.7.80	tcp
log	accept	www-http	192.168.1.60	209.133.65.45	tcp
log	accept	www-http	192.168.1.60	66.111.250.58	tcp
log	accept	www-http	192.168.1.60	66.111.250.58	tcp
log	accept	www-http	192.168.1.60	209.133.65.45	tcp
log	accept	www-http	192.168.1.60	209.15.178.121	tcp
log	accept	www-http	192.168.1.60	209.15.178.121	tcp

The issue at hand is the method that stateful inspection firewalls use to log packets that meet firewall ruleset criteria. In the Checkpoint case, the firewall has a proprietary log viewer that allows the selection of different criteria such as source and destination IP addresses, service port, protocol, and other relevant log data. The usefulness of this information becomes questionable when one is attempting to efficiently gain information about user web habits. By default, URL information is not logged in the Checkpoint logs. There is a feature to resolve IP addresses to names, but this option can seriously limit performance of the log viewer and its functionality is quite limited as the screenshot below shows:

Type	Action	Service	Source	Destination	Proto.
log	accept	www-http	192.168.1.60	209.133.65.45	tcp
log	accept	www-http	192.168.1.60	chanrr1.real.com	tcp
log	accept	www-http	192.168.1.60	209.133.65.45	tcp
log	accept	www-http	192.168.1.60	66.111.250.58	tcp
log	accept	www-http	192.168.1.60	66.111.250.58	tcp
log	accept	www-http	192.168.1.60	209.133.65.45	tcp
log	accept	www-http	192.168.1.60	209.15.178.121	tcp
log	accept	www-http	192.168.1.60	209.15.178.121	tcp

This is not intended to dismiss the log functionality with a Checkpoint Firewall-1 product or stateful inspection firewalls, but to simply note that to utilize this product to gain individual user web access habits requires several changes and/or additions to the firewall. Such changes may involve setting up user accounts on the firewall, setting up the proxy functionality of the product, or incorporating a third party product to obtain this information for a fee. In many cases some or all of these requirements are effective, but

still require manual review of the logs via a proprietary log viewer, or a process to export the logs into viewable format. In the Checkpoint Firewall-1 case, most of these changes require a firewall outage and many companies utilize the corporate firewall for much more than general purpose, employee computing. As a result other operations could be impacted to make such changes.

## The Open Source Solution

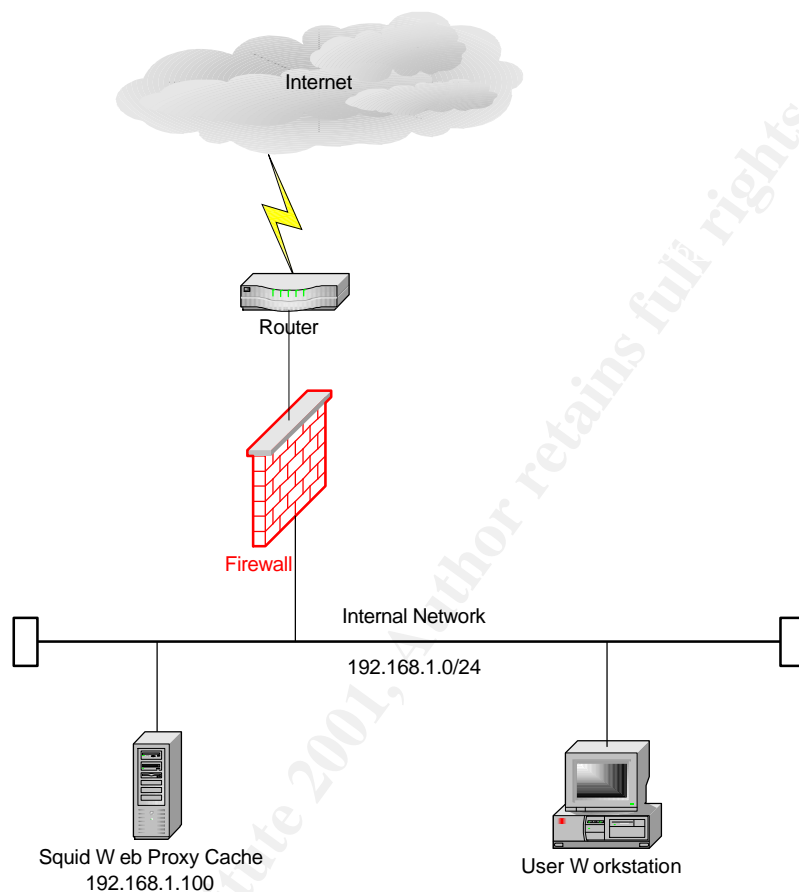
This paper is intended to get network and security administrators thinking about using an open source proxy server to provide a solution to the problem of monitoring user web habits. With a little time and effort a Squid Web Proxy Cache with other freely available tools running on Linux can provide an organization with the resource necessary to achieve the intended goal. It goes without saying that other tools can be purchased to achieve this goal. The Cacheflow Accelerator and iPlanet Proxy are examples of products which offer enterprise class proxy solutions that meet the needs and requirements of organizations. Each of these products has various attributes and features that may make the open source solution less appealing for the enterprise, but it is always good practice to see alternatives to fee based products and support. This solution is intended to be built and tested in a lab environment and put in place without ever having to take a production firewall offline.

## The Design

The design of this proxy solution is geared around a single Linux server, with Apache and openSSL compiled and running with the following applications<sup>3</sup>.

- Squid Web Cache Proxy (<http://www.squid-cache.org>)
- Apache httpd utilizing the htpasswd (username/password) utility (<http://httpd.apache.org/docs/programs/htpasswd.html>)
- SARG (Squid Analysis Report Generator) (<http://web.onda.com.br/orso/sarg.htm>)

The design in this case requires the Squid server be built with a single network interface card and be installed onto a network relatively close to both the web users and the firewall, on the internal side of the network. This design assumes network address translation is performed at the firewall. See the following network diagram:



## Building the Squid Web Cache Proxy Server

To install Squid:

1. Download the Squid rpm file (and associated dependencies) from any rpm repository (i.e. rpmfind.net) (recommend the latest STABLE release for the Linux distribution utilized):

```
squid-2.4.STABLE1-6.i386.rpm
```

2. Save the file to a temporary directory (i.e. /var/tmp)
3. Execute and install the Squid cache with the following command:

```
rpm --install --verbose --hash squid-2.4.STABLE1-6.i386.rpm
```

(this command is abbreviated using `'rpm -ivh squid-2.4.STABLE1-6.i386.rpm'`)

4. The installation builds several directories including:

```
/etc/squid – squid configuration files  
/etc/rc.d/init.d/squid – squid daemon  
/var/log/squid – squid logs
```

5. Change directory to `/etc/squid` and use a text editor to edit the file `squid.conf`. This is the main configuration file and must be altered for the squid daemon to respond to requests. Squid is finely documented at <http://squid.visolve.com/squid24s1/contents.htm>. The minimum configuration options for the purposes of this document will be addressed here:

*By default Squid runs on tcp port 3128. For the sake of this document the default will be changed to tcp port 8080 under the 'TAG: http\_port'.*

### **NETWORK OPTIONS**

```
# TAG: http_port  
http_port 8080 – port the Squid daemon will listen on.  
icp_port 0 – disable icp for this document  
htcp_port 0 – disable htcp for this document
```

*To setup required proxy authentication, make the following bolded, blue changes<sup>4</sup>. The 'TAG: authenticate\_program' determines what authenticator will*

be used to authenticate users. The rpm installs the ncsa\_auth utility in a different directory than that of the example.

#### OPTIONS FOR EXTERNAL SUPPORT PROGRAMS

```
# TAG: authenticate_program
#Default:
authenticate_program /usr/lib/squid/ncsa_auth
/etc/squid/squid_passwd
```

Squid access control lists (TAG: acl) are read from the top down (similar to Cisco IOS ACLs); however, they do not account for an implicit “deny all”. By default the Squid configuration file denies all access via the ‘http\_access deny all’ setting. In order to change this, ACLs for individual access must be added before the explicit “deny all” setting. It is highly recommend that access to the Squid engine be controlled in these settings.

A sample network ACL with its corresponding http\_access entry is shown for additional information. The ACL ‘proxy\_auth REQUIRED’ entry forces user authentication as defined in the ‘authenticate program’ tag and it’s placement in at the top of the access control list.

#### ACCESS CONTROLS

```
# TAG: acl
acl user_auth proxy_auth REQUIRED
#acl net_192.168.1.0 src 192.168.1.0/255.255.255.0
acl all src 0.0.0.0/0.0.0.0
http_access allow manager localhost
http_access deny manager
http_access deny !Safe_ports
http_access deny CONNECT !SSL_ports

# INSERT YOUR OWN RULE(S) HERE TO ALLOW ACCESS FROM YOUR
CLIENTS
#
# And finally deny all other access to this proxy
http_access allow localhost
http_access allow user_auth
#http_access allow net_192.168.1.0
http_access deny all
```



*In order to store a certain number of previously rotated logs, alter the 'TAG'logfile\_rotate' entry from the default 0 to a preferred value. The value is the number of old log files Squid will store in the /var/log/squid directory.*

#### **MICELLANEOUS**

```
# TAG: logfile_rotate
logfile_rotate 10
```

6. Start Squid with the following command `\usr/sbin/squid -d 1'`. This will enable debugging for review. It is a good idea to have Squid start up on default. Use `\sbin/chkconfig'` or the command `\usr/sbin/ntsysv'` to have Squid startup at boot time. Note that `\usr/sbin/ntsysv'` must be changed per runlevel. By default `\usr/sbin/ntsysv'` with no switches configures the daemon for the current run level only.

Examples:

```
/sbin/chkconfig -levels 2345 squid on
```

(turns daemon on for runlevels 2,3,4,5)

```
/usr/sbin/ntsysv -level 2345
```

(spawns graphical interface to change daemon startup for specified runlevels 2,3,4,5)

Use `\sbin/chkconfig -list squid'` squid to verify what runlevels squid is set to startup in.

## **Adding Users**

In order to define users and user passwords, utilize the 'htpasswd' command compiled with Apache. Htpasswd is a simple utility for creating and storing usernames and encrypted passwords, in a single file, on the system. The initial command to create the password file and populate it with the first user is as follows:

```
/usr/bin/htpasswd -c [password_file] [username]
```

```
/usr/bin/htpasswd -c /etc/squid/squid_passwd jdoe
```

```
New password:
```

```
Re-type new password:
```

```
Adding password for user jdoe
```

Subsequent passwords can be created. Ensure that the '-c' switch is removed as this switch will overwrite the existing file with a new file. For example:

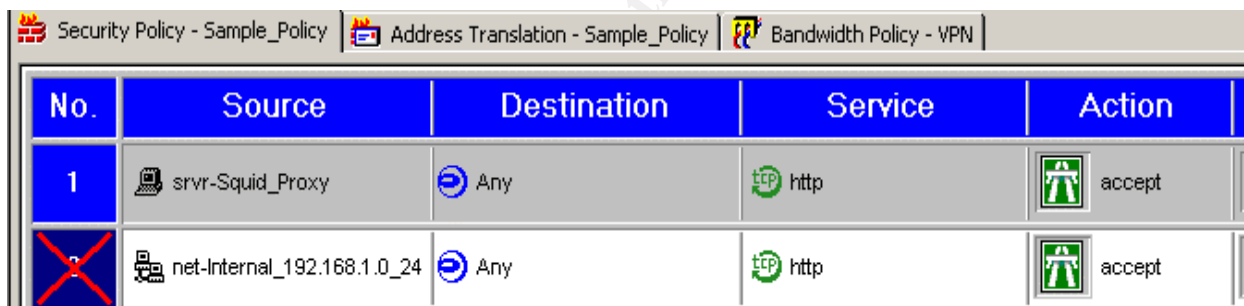
```
/usr/bin/htpasswd /etc/squid/squid_passwd jdgregg
New password:
Re-type new password:
Adding password for user jdgregg
```

In order for new usernames to take effect squid must be reconfigured with the following command:

```
/usr/sbin/squid -k reconfigure
```

## Firewall Changes

If a stateful inspection firewall is utilized, some changes to only allow the Squid server to access the Internet for http services will need to be made. Since the Squid box is proxying client sessions to the Internet, eliminating client access directly through the firewall for http access is recommended. An altered Checkpoint FW-1 rule may look as follows:

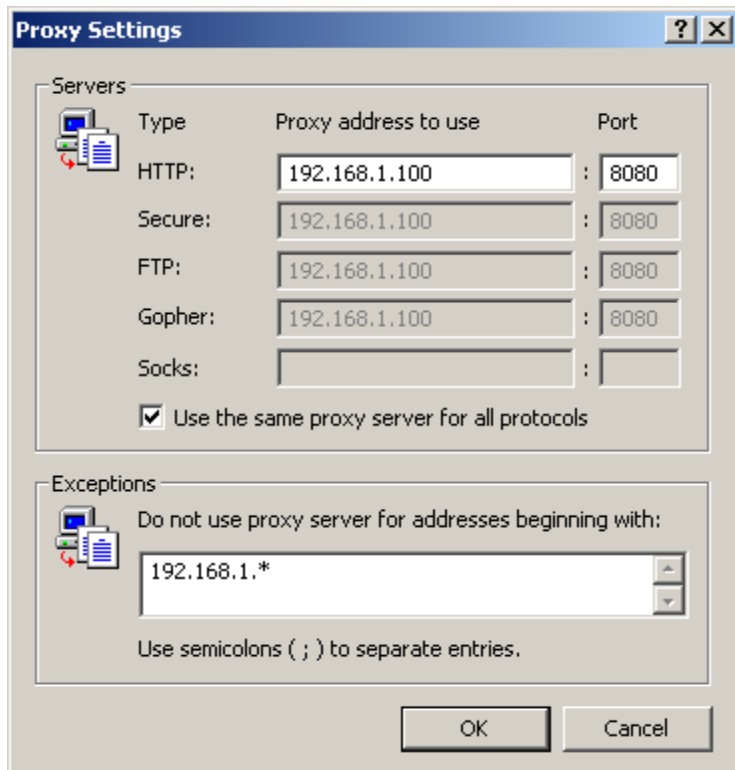


No.	Source	Destination	Service	Action
1	srvr-Squid_Proxy	Any	http	accept
	net-Internal_192.168.1.0_24	Any	http	accept

(Note: the large red "X" indicates old http access rule is disabled and once a policy is installed will no longer pass traffic)

## Browser Settings

Browser settings need to be altered for all http and browser ftp traffic to redirect to the Squid server. The settings for individual browsers can typically be found in the network options.



© SANS Institute 2001, Author retains full rights

## Browser Authentication Interface

With `'proxy_auth REQUIRED'` set in `squid.conf`, the user interface for authenticating is prompted upon opening the browser and attempting to access a site off of the local network.



Users attempting to authenticate without a valid user account will receive an error message from the Squid server. These error messages can be altered to meet the needs of an individual organization. Squid error messages can be found in `/etc/squid/errors`

## User Monitoring Tool

In order to begin generating reports on user web access, a report generator will need to be installed. Many Squid reporting tools are available, but SARG (Squid Analysis Report Generator) was chosen due to its clean reporting style, completeness, and ease of use. The `'/var/log/squid/access.log'` is the log that stores individual user web access and is most important for the purposes of this document.

1. Install the SARG rpm with the following command;

```
rpm -ivh sarg-1.1-1.i386.rpm
```

2. Change to the configuration directory `'/etc/sarg'`. Edit the `sarg.conf` file.
3. The defaults can be accepted in most cases except where changes meet the specific requirements of an organization. A change to the default directory where

squid reports are kept may be necessary to match the default html directory of the Apache server.

```
# TAG:  output_dir
#      Where is the reports will be stored.
#      sarg -o dir
#
#output_dir /usr/local/etc/httpd/htdocs/squid-reports
output_dir /var/www/html/squid-reports
```

4. To rotate the squid logs use the command:

```
/usr/sbin/squid -k rotate
```

This will generate a new log and rename the previous access.log (i.e. access.log.0)

5. The command to run a set of reports for a given access.log file is as follows.

```
/usr/sbin/sarg -l /var/log/squid/access.log.0
```

```
SARG: Successful report generated on /var/www/html/squid-
reports/2001Nov19-2001Nov19
```

The '-l' switch is the location of the input log that SARG will run a report against.

6. Once the reports are generated they are stored in the output directory specified in the `sarg.conf` file. The `httpd` and `sarg` directories need to be the same if remote access of the reports is desired<sup>5</sup>. A cron entry that rotates the logs then runs a `sarg` report against those logs is recommended.

© SANS Institute 2001. Author retains full rights.

## Analyzing SARG Reports

The report structure is simple and contains quite a bit of useful information. The initial page shows the dated report:

Squid User Access Report				
FILE/PERIOD	CREATION DATE	USERS	BYTES	AVERAGE
<a href="#">2001Nov19-2001Nov19</a>	Mon Nov 19 16:29:38 CST 2001	4	4,647,624	1,161,906

Generated by [sarg 1.1.15Mar2001](#) on Nov19 2001 16:29

The report is simply organized with useful links. The primary page gives a concise listing of users who have utilized the proxy server as well as the time a user spent accessing the proxy and other relevant data.

Squid User Access Report								
Period: 2001Nov19-2001Nov19								
Sort: BYTES, reverse								
Topuser Report								
<a href="#">Topsites</a> Report								
<a href="#">Sites &amp; Users</a> Report								
NUM	USERID	CONNECT	BYTES	%BYTESIN-CACHE	OUTUSED	TIMEMILISEC	%TIME	
1	<a href="#">date/time</a> <a href="#">jdgregg</a>	7202	1,596,696	46.47%	7.14% 92.86%	00:04:18	258.724	43.20%
2	<a href="#">date/time</a> <a href="#">jdoe</a>	3991	3,724,438	29.53%	1.93% 98.07%	00:03:33	213.194	35.59%
3	<a href="#">date/time</a> <a href="#">bsmith</a>	3841	1,111,382	23.91%	1.73% 98.27%	00:02:07	127.031	21.21%
4	<a href="#">date/time</a> <a href="#">192.168.1.11</a>	3	4,108	0.09%	100.00% 0.00%	00:00:00	6	0.00%
<b>TOTAL</b>			<b>1,506,647,624</b>		<b>4.39% 95.61%</b>	<b>00:09:58</b>	<b>598.955</b>	
<b>AVERAGE</b>			<b>376,161,906</b>			<b>00:02:29</b>	<b>149.738</b>	

A drilldown into the html links in the report of a specific user gives an administrator an excellent view of web habits. For example a drilldown of user *jdgregg* shows web habits focused around sports themes that may or may not be condoned while at work:

Squid User Access Report								
Period: 2001Nov19-2001Nov19								
User: jdgregg								
Sort: BYTES, reverse								
User Report								
ACCESSED SITE	CONNECT	BYTES	%BYTES	IN-CACHE	OUT	USED TIME	MILISEC	%TIME
<a href="http://www.clevelandbrowns.com">www.clevelandbrowns.com</a>	63	308.976	14.31%	0.00%	100.00%	00:00:26	26.174	10.12%
<a href="http://www.miamidolphins.com">www.miamidolphins.com</a>	40	214.990	9.95%	0.00%	100.00%	00:00:26	26.836	10.37%
<a href="http://www.buffalobills.com">www.buffalobills.com</a>	111	204.621	9.47%	54.97%	45.03%	00:00:07	7.135	2.76%
<a href="http://www.bengals.com">www.bengals.com</a>	48	175.482	8.13%	0.12%	99.88%	00:00:17	17.322	6.70%
<a href="http://www.colts.com">www.colts.com</a>	36	173.016	8.01%	0.12%	99.88%	00:00:15	15.702	6.07%
<a href="http://www.jaguars.com">www.jaguars.com</a>	45	159.508	7.39%	0.00%	100.00%	00:00:14	14.424	5.58%
<a href="http://games.espn.go.com">games.espn.go.com</a>	57	141.778	6.56%	0.00%	100.00%	00:00:21	21.160	8.18%
<a href="http://www.newyorkjets.com">www.newyorkjets.com</a>	23	138.804	6.43%	0.00%	100.00%	00:00:12	12.139	4.69%
<a href="http://sports.espn.go.com">sports.espn.go.com</a>	2	79.366	3.67%	0.00%	100.00%	00:00:02	2.818	1.09%
<a href="http://www.burstnet.com">www.burstnet.com</a>	16	59.683	2.76%	1.40%	98.60%	00:00:04	4.199	1.62%
<a href="http://football.weblogs.com">football.weblogs.com</a>	1	54.163	2.51%	0.00%	100.00%	00:00:15	15.552	6.01%
<a href="http://www.itradecards.com">www.itradecards.com</a>	1	41.245	1.91%	0.00%	100.00%	00:00:01	1.408	0.54%
<a href="http://www.packers.com">www.packers.com</a>	2	38.699	1.79%	0.00%	100.00%	00:00:04	4.500	1.74%
<a href="http://www.tackls.com">www.tackls.com</a>	5	38.249	1.77%	2.19%	97.81%	00:00:01	1.219	0.47%
<a href="http://www.geocities.com">www.geocities.com</a>	48	37.256	1.73%	25.84%	74.16%	00:00:01	1.758	0.68%
<a href="http://bannerads.miamidolphins.com">bannerads.miamidolphins.com</a>	3	36.942	1.71%	0.00%	100.00%	00:00:04	4.902	1.89%
<a href="http://www.e-jocks.com">www.e-jocks.com</a>	3	33.378	1.55%	0.00%	100.00%	00:00:02	2.175	0.84%
<a href="http://www.youthfantasyfootball.com">www.youthfantasyfootball.com</a>	1	25.885	1.20%	0.00%	100.00%	00:00:01	1.030	0.40%
<a href="http://static.userland.com">static.userland.com</a>	14	24.773	1.15%	0.00%	100.00%	00:00:03	3.929	1.52%
<a href="http://downloads.weblogger.com">downloads.weblogger.com</a>	17	17.884	0.83%	0.00%	100.00%	00:00:05	5.981	2.31%
<a href="http://ad.espn.starwave.com">ad.espn.starwave.com</a>	4	16.252	0.75%	0.00%	100.00%	00:00:01	1.520	0.59%
<a href="http://www2.miamidolphins.com">www2.miamidolphins.com</a>	13	16.100	0.75%	0.00%	100.00%	00:00:06	6.060	2.34%
<a href="http://www.krause.com">www.krause.com</a>	30	15.737	0.73%	100.00%	0.00%	00:00:00	301	0.12%
<a href="http://espn.go.com">espn.go.com</a>	6	13.543	0.63%	0.00%	100.00%	00:00:04	4.469	1.73%
<a href="http://proshop.bengals.com">proshop.bengals.com</a>	1	11.575	0.54%	0.00%	100.00%	00:00:00	973	0.38%
<a href="http://www.zoneswap.com">www.zoneswap.com</a>	3	9.281	0.43%	0.00%	100.00%	00:00:17	17.743	6.86%
<a href="http://www.nfl.com">www.nfl.com</a>	6	7.593	0.35%	2.77%	97.23%	00:00:02	2.882	1.11%
<a href="http://www.toughguy.com">www.toughguy.com</a>	35	7.330	0.34%	100.00%	0.00%	00:00:00	320	0.12%
<a href="http://images.nfl.com">images.nfl.com</a>	4	6.745	0.31%	9.30%	90.70%	00:00:00	360	0.14%
<a href="http://stats.hitbox.com">stats.hitbox.com</a>	1	5.269	0.24%	0.00%	100.00%	00:00:01	1.320	0.51%



Further drilldowns show other interesting and questionable web habits. User *jdoe* in particular is utilizing work hours accessing sites that could be deemed highly suspect and in some cases illegal or dangerous:

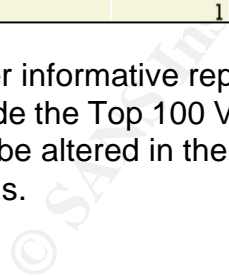
Squid User Access Report								
Period: 2001Nov19-2001Nov19								
User: jdoe								
Sort: BYTES, reverse								
User Report								
ACCESSED SITE	CONNECT	BYTES	%BYTES	IN-CACHE	OUT	USED TIME	MILISEC	%TIME
<a href="http://www.casino.net">www.casino.net</a>	36	166.201	12.11%	0.00%	100.00%	00:00:20	20.938	9.82%
<a href="http://www2.warnerbros.com">www2.warnerbros.com</a>	18	151.887	11.07%	0.00%	100.00%	00:00:13	13.132	6.16%
<a href="http://www.penthouse.com">www.penthouse.com</a>	4	140.518	10.24%	0.00%	100.00%	00:00:04	4.871	2.28%
<a href="http://www.2600.com">www.2600.com</a>	21	126.263	9.20%	0.00%	100.00%	00:00:11	11.934	5.60%
<a href="http://genres.mp3.com">genres.mp3.com</a>	3	125.272	9.13%	0.00%	100.00%	00:00:04	4.600	2.16%
<a href="http://grfx.mp3.com">grfx.mp3.com</a>	11	68.750	5.01%	0.00%	100.00%	00:00:11	11.833	5.55%
<a href="http://play.games.com">play.games.com</a>	69	62.713	4.57%	22.90%	77.10%	00:00:02	2.141	1.00%
<a href="http://download.cnet.com">download.cnet.com</a>	10	62.376	4.54%	0.00%	100.00%	00:00:09	9.778	4.59%
<a href="http://www.astalavista.com">www.astalavista.com</a>	13	57.587	4.20%	3.78%	96.22%	00:00:04	4.097	1.92%
<a href="http://www.mp3.com">www.mp3.com</a>	12	51.377	3.74%	0.00%	100.00%	00:00:05	5.678	2.66%
<a href="http://img.napster.com">img.napster.com</a>	38	50.403	3.67%	0.00%	100.00%	00:00:15	15.517	7.28%
<a href="http://images.mp3.com">images.mp3.com</a>	16	30.240	2.20%	0.00%	100.00%	00:00:14	14.734	6.91%
<a href="http://a.r.tv.com">a.r.tv.com</a>	9	28.600	2.08%	0.00%	100.00%	00:00:02	2.645	1.24%
<a href="http://www.icq.cracks.ru">www.icq.cracks.ru</a>	13	27.819	2.03%	0.00%	100.00%	00:00:27	27.121	12.72%
<a href="http://www.napster.com">www.napster.com</a>	4	27.603	2.01%	0.00%	100.00%	00:00:02	2.366	1.11%
<a href="http://ad.pbs.bb.ru">ad.pbs.bb.ru</a>	6	27.280	1.99%	0.00%	100.00%	00:00:04	4.127	1.94%
<a href="http://login.mp3.com:443">login.mp3.com:443</a>	6	20.928	1.52%	0.00%	100.00%	00:00:04	4.194	1.97%
<a href="http://www.cracks.ru">www.cracks.ru</a>	11	17.151	1.25%	5.06%	94.94%	00:00:19	19.896	9.33%
<a href="http://lc2.law5.hotmail.passport.com">lc2.law5.hotmail.passport.com</a>	2	15.930	1.16%	0.00%	100.00%	00:00:01	1.830	0.86%
<a href="http://www.anonymizer.org">www.anonymizer.org</a>	35	15.634	1.14%	45.89%	54.11%	00:00:00	698	0.33%
<a href="http://ng3.ads.warnerbros.com">ng3.ads.warnerbros.com</a>	4	13.553	0.99%	0.00%	100.00%	00:00:02	2.388	1.12%
<a href="http://217.170.71.61">217.170.71.61</a>	1	12.266	0.89%	0.00%	100.00%	00:00:00	976	0.46%
<a href="http://toolbar.netscape.com">toolbar.netscape.com</a>	9	10.504	0.77%	12.05%	87.95%	00:00:01	1.334	0.63%
<a href="http://m.doubleclick.net">m.doubleclick.net</a>	1	9.692	0.71%	0.00%	100.00%	00:00:00	685	0.32%
<a href="http://www.playboy.com">www.playboy.com</a>	2	6.958	0.51%	3.31%	96.69%	00:00:00	487	0.23%
<a href="http://mozg.hobot.ru">mozg.hobot.ru</a>	7	6.814	0.50%	3.08%	96.92%	00:00:06	6.061	2.84%
<a href="http://216.32.242.251">216.32.242.251</a>	5	6.620	0.48%	0.00%	100.00%	00:00:02	2.031	0.95%
<a href="http://includes.mp3.com">includes.mp3.com</a>	3	6.542	0.48%	0.00%	100.00%	00:00:03	3.905	1.83%
<a href="http://www.wb.com">www.wb.com</a>	2	5.892	0.43%	0.00%	100.00%	00:00:00	652	0.31%
<a href="http://ul18.35.spylog.com">ul18.35.spylog.com</a>	2	3.711	0.27%	0.00%	100.00%	00:00:01	1.092	0.51%



A look at user **bsmith** shows what may be deemed acceptable, expected web habits during working hours:

<b>Squid User Access Report</b>								
Period: 2001Nov19-2001Nov19								
User: bsmith								
Sort: BYTES, reverse								
User Report								
ACCESSED SITE	CONNECT	BYTES	%BYTES	IN-CACHE	OUT	USED TIME	MILISEC	%TIME
<a href="http://www.netgear.com">www.netgear.com</a>	65	185.166	16.66%	0.00%	100.00%	00:00:21	21.806	17.17%
<a href="http://www.netiq.com">www.netiq.com</a>	55	163.842	14.74%	0.61%	99.39%	00:00:19	19.551	15.39%
<a href="http://www.foundrynetworks.com">www.foundrynetworks.com</a>	29	147.318	13.26%	2.99%	97.01%	00:00:05	5.299	4.17%
<a href="http://www.redhat.com">www.redhat.com</a>	41	128.577	11.57%	0.00%	100.00%	00:00:10	10.210	8.04%
<a href="http://www.be.com">www.be.com</a>	32	114.321	10.29%	0.00%	100.00%	00:00:11	11.305	8.90%
<a href="http://www.snort.org">www.snort.org</a>	6	103.600	9.32%	0.00%	100.00%	00:00:12	12.763	10.05%
<a href="http://www.rsa.com">www.rsa.com</a>	43	89.393	8.04%	0.00%	100.00%	00:00:14	14.767	11.62%
<a href="http://www.ibm.com">www.ibm.com</a>	21	69.405	6.24%	0.00%	100.00%	00:00:06	6.245	4.92%
<a href="http://www.microsoft.com">www.microsoft.com</a>	21	55.833	5.02%	4.28%	95.72%	00:00:05	5.490	4.32%
<a href="http://www.cisco.com">www.cisco.com</a>	24	24.702	2.22%	19.74%	80.26%	00:00:00	837	0.66%
<a href="http://m.doubleclick.net">m.doubleclick.net</a>	1	7.405	0.67%	0.00%	100.00%	00:00:00	474	0.37%
<a href="http://cco-sj-l.cisco.com">cco-sj-l.cisco.com</a>	1	5.825	0.52%	0.00%	100.00%	00:00:00	357	0.28%
<a href="http://www.sans.org">www.sans.org</a>	25	5.611	0.50%	90.25%	9.75%	00:00:00	582	0.46%
<a href="http://www.redhat.com:443">www.redhat.com:443</a>	2	3.497	0.31%	0.00%	100.00%	00:00:02	2.783	2.19%
<a href="http://stats.klsoft.com">stats.klsoft.com</a>	3	1.834	0.17%	0.00%	100.00%	00:00:01	1.016	0.80%
<a href="http://www.rpmfind.net.com">www.rpmfind.net.com</a>	1	1.131	0.10%	0.00%	100.00%	00:00:00	109	0.09%
<a href="http://www.rpmfind.net">www.rpmfind.net</a>	5	1.056	0.10%	100.00%	0.00%	00:00:00	106	0.08%
<a href="http://statse.webtrendslive.com">statse.webtrendslive.com</a>	1	782	0.07%	0.00%	100.00%	00:00:00	841	0.66%
<a href="http://ng-prod1.cisco.com">ng-prod1.cisco.com</a>	2	581	0.05%	0.00%	100.00%	00:00:00	345	0.27%
<a href="http://c.microsoft.com">c.microsoft.com</a>	1	423	0.04%	0.00%	100.00%	00:00:00	336	0.26%
<a href="http://www.foundrynet.com">www.foundrynet.com</a>	2	420	0.04%	100.00%	0.00%	00:00:00	9	0.01%
<a href="http://www.dotmumble.com">www.dotmumble.com</a>	1	399	0.04%	0.00%	100.00%	00:00:00	345	0.27%
<a href="http://ad.doubleclick.net">ad.doubleclick.net</a>	1	261	0.02%	0.00%	100.00%	00:00:00	235	0.18%
<a href="http://sans.org">sans.org</a>	1	0	0.00%	0.00%	0.00%	00:00:11	11.220	8.83%

Finally, other informative reports are generated and accessible from the initial page. These include the Top 100 Visited Web sites and a Web Site to User correlation. These reports can be altered in the `sarg.conf` file to meet the needs of individual organizations.



## Conclusion

With the a little time and effort, a system administrator can utilize a Squid Proxy solution to monitor user web habits for little to no cost outside of the hardware expenses. This document only touches on some of the functionality that Squid and the other open source software discussed in this document can provide for an organization. Over time the solution presented here could grow into a geographically diverse design that provides transparent caching, third-party authentication, and a multi-tier caching hierarchy. The beauty is Squid, and the those that support it, will continue to meet the needs of the community that presents these challenges.

© SANS Institute 2001, Author retains full rights

## Sources

Apache HTTP Server Version 2.0. "Manual Page: htpasswd". No date. URL: <http://httpd.apache.org/docs-2.0/programs/htpasswd.html> (20 Nov 2001).

Cacheflow Corporation. "CacheFlow Internet Caching Appliances: Next Generation Proxy Server Solution." October, 1999. URL: [http://www.cacheflow.com/files/whitepapers/wp\\_proxy\\_server.pdf](http://www.cacheflow.com/files/whitepapers/wp_proxy_server.pdf) (20 Nov. 2001).

Computer Emergency Response Team (CERT). "Design the Firewall System". Security Improvement Modules. July, 1999. URL: <http://www.cert.org/security-improvement/practices/p053.html> (20 Nov 2001).

iPlanet E-Commerce Solutions. "iPlanet Proxy Server: Overview". 2001, URL: [http://www.iplanet.com/products/iplanet\\_proxy/home\\_2\\_1\\_1ae.html](http://www.iplanet.com/products/iplanet_proxy/home_2_1_1ae.html) (20 Nov 2001).

Pearson, Oskar. "Squid: A User's Guide". Version 0.1. 2000. URL: <http://squid-docs.sourceforge.net/latest/html/book1.htm> (20 Nov 2001).

SARG (Squid Analysis Report Generator). "README". Stable Release 1.1.1. April, 2001. URL: <http://web.onda.com.br/orso/sarg.README.txt> (20 Nov 2001).

Visolve.com. "Squid 2.4 Stable1 Configuration Manual". No Date. URL: <http://squid.visolve.com/squid24s1/contents.htm> (20 Nov 2001).

© SANS Institute 2001

## References

1 Both proxy based firewalls as well as stateful inspection platforms have their positives and negatives and such a discussion is outside the scope of this paper. For the purposes of this document, CERT's definition of stateful inspection is utilized to differentiate between vendor discrepancies <http://www.cert.org/security-improvement/practices/p053.html>

2 This document will use Checkpoint's Firewall-1 product for discussing the stateful inspection URL logging limitations

3 Installing and hardening a Linux server falls outside of the scope of this document. For the purposes of this document `squid-2.4.STABLE1-6.i386` is used on a Red Hat Linux 7.2 distribution (Linux Kernel 2.4.7-10). Other distributions may require different squid daemons and setup of these Squid daemons could be different than the Squid 2.4 version described in this document. Apache version 1.3.20 was utilized for the purposes of this document.

4 The Squid rpm file automatically compiles the ncsa authenticator. Various other methods of authentication can be compiled and setup with Squid. More information can be found at <http://www.squid-cache.org/related-software.html>.

5 It is necessary to control access to these files as they may contain sensitive information.

© SANS Institute 2001, Author retains full rights



# Upcoming SANS Training

[Click here to view a list of all SANS Courses](#)

SANS Bangalore 2019	Bangalore, IN	Nov 25, 2019 - Nov 30, 2019	Live Event
SANS Cyber Threat Summit 2019	London, GB	Nov 25, 2019 - Nov 26, 2019	Live Event
SANS Nashville 2019	Nashville, TNUS	Dec 02, 2019 - Dec 07, 2019	Live Event
SANS San Francisco Winter 2019	San Francisco, CAUS	Dec 02, 2019 - Dec 07, 2019	Live Event
SANS Security Operations London 2019	London, GB	Dec 02, 2019 - Dec 07, 2019	Live Event
SANS Paris December 2019	Paris, FR	Dec 02, 2019 - Dec 07, 2019	Live Event
SANS Frankfurt December 2019	Frankfurt, DE	Dec 09, 2019 - Dec 14, 2019	Live Event
SANS Cyber Defense Initiative 2019	Washington, DCUS	Dec 10, 2019 - Dec 17, 2019	Live Event
SANS Austin Winter 2020	Austin, TXUS	Jan 06, 2020 - Jan 11, 2020	Live Event
SANS Threat Hunting & IR Europe Summit & Training 2020	London, GB	Jan 13, 2020 - Jan 19, 2020	Live Event
SANS Miami 2020	Miami, FLUS	Jan 13, 2020 - Jan 18, 2020	Live Event
Cyber Threat Intelligence Summit & Training 2020	Arlington, VAUS	Jan 20, 2020 - Jan 27, 2020	Live Event
SANS Tokyo January 2020	Tokyo, JP	Jan 20, 2020 - Jan 25, 2020	Live Event
SANS Amsterdam January 2020	Amsterdam, NL	Jan 20, 2020 - Jan 25, 2020	Live Event
SANS Anaheim 2020	Anaheim, CAUS	Jan 20, 2020 - Jan 25, 2020	Live Event
MGT521 Beta Two 2020	San Diego, CAUS	Jan 22, 2020 - Jan 23, 2020	Live Event
SANS Las Vegas 2020	Las Vegas, NVUS	Jan 27, 2020 - Feb 01, 2020	Live Event
SANS Vienna January 2020	Vienna, AT	Jan 27, 2020 - Feb 01, 2020	Live Event
SANS San Francisco East Bay 2020	Emeryville, CAUS	Jan 27, 2020 - Feb 01, 2020	Live Event
SANS Security East 2020	New Orleans, LAUS	Feb 01, 2020 - Feb 08, 2020	Live Event
SANS London February 2020	London, GB	Feb 10, 2020 - Feb 15, 2020	Live Event
SANS Northern VA - Fairfax 2020	Fairfax, VAUS	Feb 10, 2020 - Feb 15, 2020	Live Event
SANS New York City Winter 2020	New York City, NYUS	Feb 10, 2020 - Feb 15, 2020	Live Event
SANS Dubai February 2020	Dubai, AE	Feb 15, 2020 - Feb 20, 2020	Live Event
SANS Brussels February 2020	Brussels, BE	Feb 17, 2020 - Feb 22, 2020	Live Event
SANS Scottsdale 2020	Scottsdale, AZUS	Feb 17, 2020 - Feb 22, 2020	Live Event
SANS San Diego 2020	San Diego, CAUS	Feb 17, 2020 - Feb 22, 2020	Live Event
Open-Source Intelligence Summit & Training 2020	Alexandria, VAUS	Feb 18, 2020 - Feb 24, 2020	Live Event
SANS Tokyo November 2019	Online.JP	Nov 25, 2019 - Nov 30, 2019	Live Event
SANS OnDemand	Books & MP3s OnlyUS	Anytime	Self Paced