UT05: Instalación y configuración de servidores proxy – 2) Complementos Squid.

Nombre: Francisco Jesús García – Uceda Díaz - Albo

Curso: 2º ASIR.

Índice

Introducción

En esta práctica aprenderemos sobre las listas negras y aplicaciones complementarias de Squid como Sarg y Calamaris. Aprenderemos sobre estos programas y como pueden realizar estadísticas mediante el archivo de log de Squid.

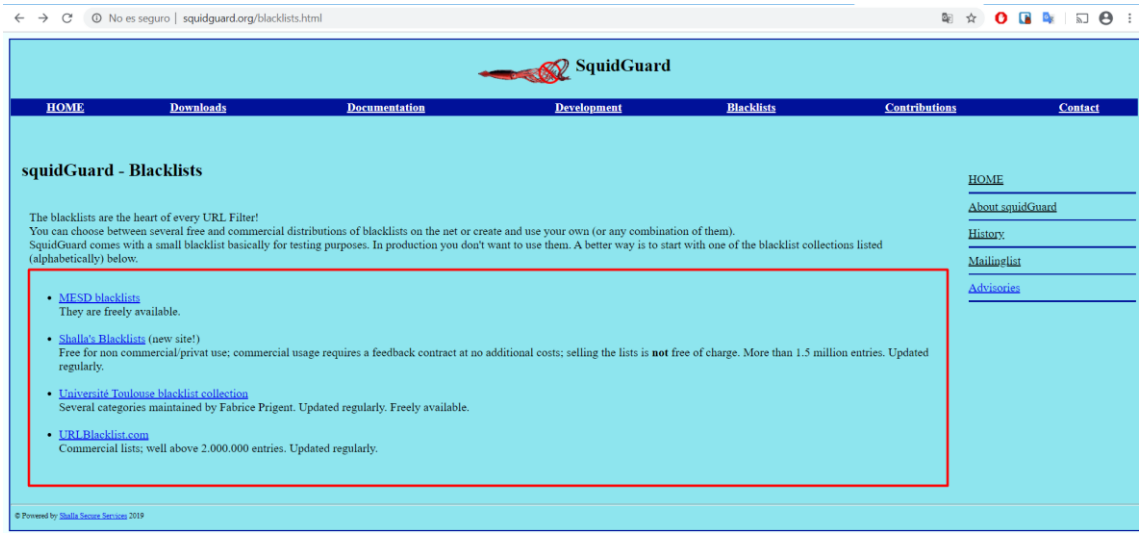
2. Complementos Squid:

a) Descarga archivo de listas negras de sitios web y anexa las mismas a la configuración de Squid. http://urlblacklist.com/ com/. Prueba dichas listas negras.

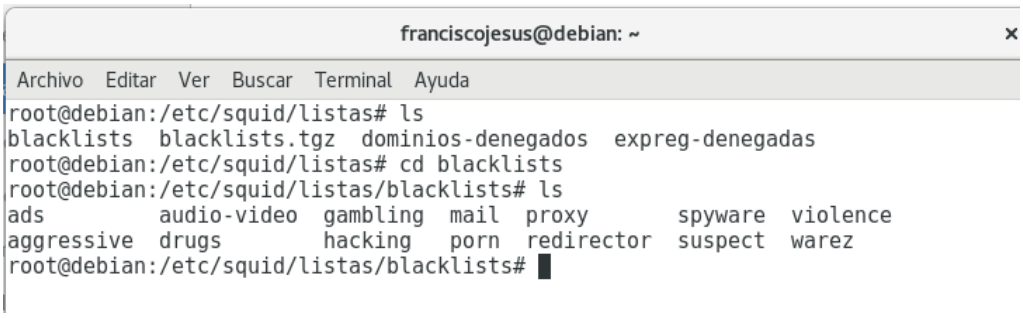
El uso de listas negras es una práctica muy común para proteger a los usuarios contra ataques, spyware, spam o contenido web indecente o potencialmente peligroso.

Las listas negras (o blacklists) son listados donde se incluyen aquellos sitios y/o direcciones IP que se desean bloquear su acceso.

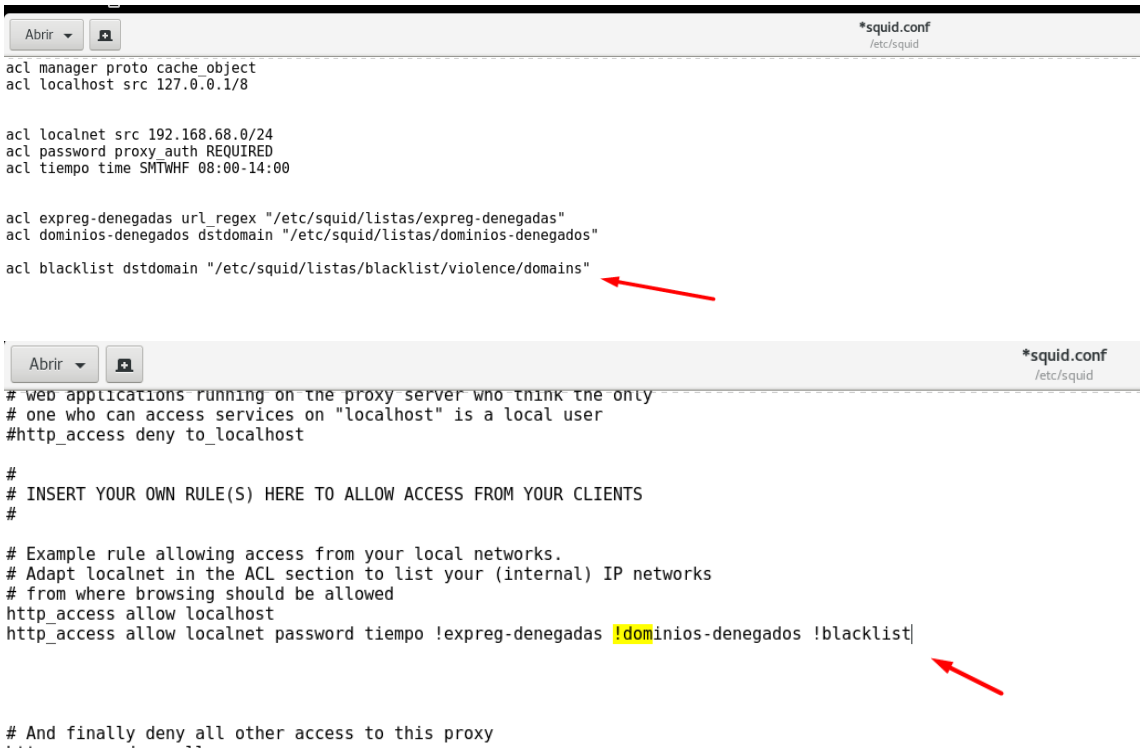
Descargamos la lista negra.



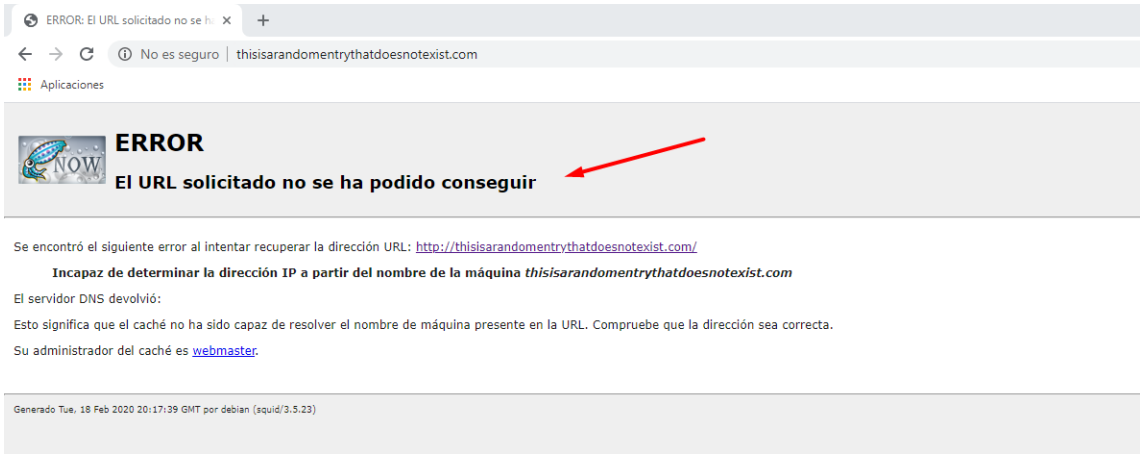
Podemos visualizar como contiene una gran cantidad de dominios.

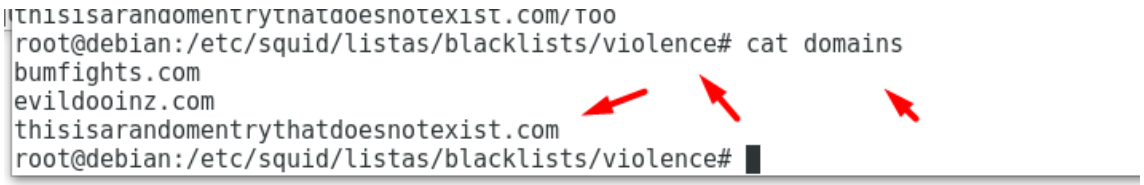


La aplicamos a nuestro servidor proxy Squid con una ACL.



Podemos probar en el cliente como funciona correctamente y bloquea los dominios de la lista.

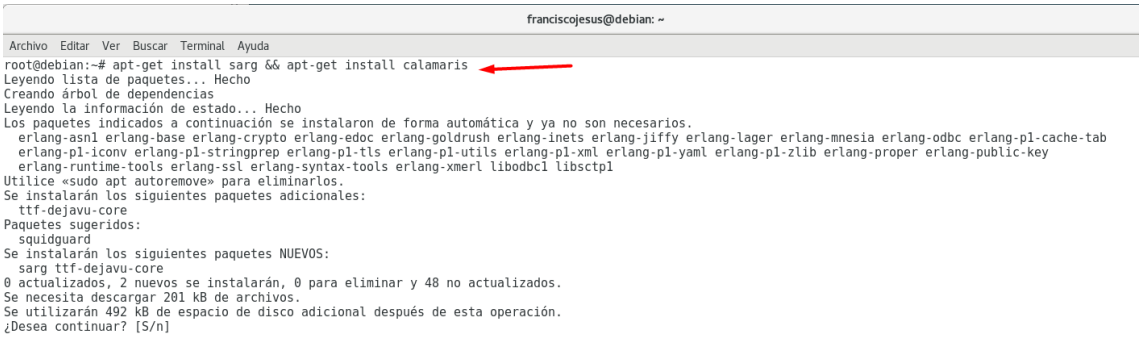




b) Utiliza las aplicaciones Sarg http://sarg.sourceforge.net/ y Calamaris http://Calamaris.Cord.de para generar informes y estadísticas de Squid a partir de los archivos log de Squid.

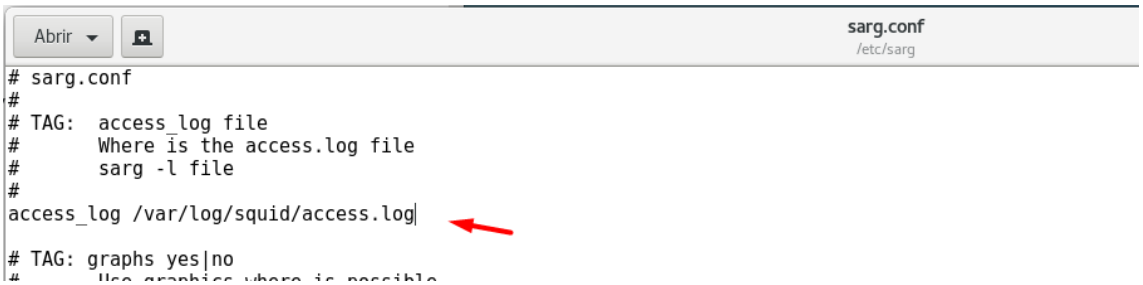
Sarg (Squid Analysis Report Generator por sus siglas en inglés) es un generador de reportes de análisis de tráfico que funciona con el Proxy Squid en Linux.

Instalamos Sarg y Calamaris.

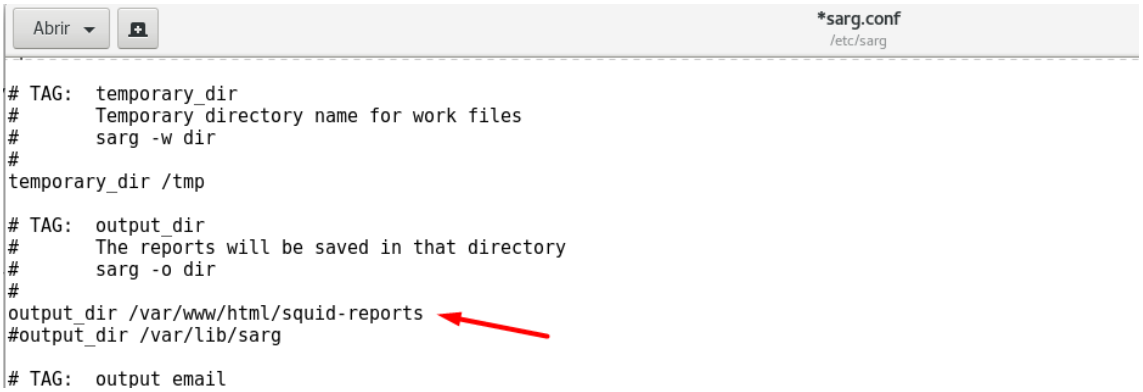


Empezaremos configurando sarg. Vamos a su archivo de configuración situado en */etc/sarg/sarg.conf*

Verificaremos que la línea de donde coge el archivo log sea */var/log/access.log*



Después configuraremos el directorio en el cual se irán guardando los informes de SARG, por medio de la directiva *output_dir*.



También podemos configurar el formato de fecha con la directiva: date_format (ya que por defecto está en notación americana). Para cambiarlo al formato europeo modificamos lo siguiente.

Abrir *sarg.conf /etc/sarg

```
#
exclude_hosts /etc/sarg/exclude_hosts

# TAG:  useragent_log file
#       useragent.log file patch to generate useragent report.
#
#useragent_log none

# TAG:  date_format
#       Date format in reports: e (European=dd/mm/yy), u (American=mm/dd/yy), w (Weekly=yy.ww)
#
date_format e

# TAG:  per_user_limit file MB
```

Configuramos para que resuelva las IP.

Abrir *sarg.conf /etc/sarg

```
#output_dir /var/lib/sarg

# TAG:  output_email
#       Email address to send the reports. If you use this tag, no html reports will be
#       sarg -e email
#
#output_email none

# TAG:  resolve_ip yes/no
#       Convert ip address to dns name
#       sarg -n
resolve_ip yes
```

Generamos un informe de Sarg con el comando ***sarg -x***

franciscojesus@debian: ~

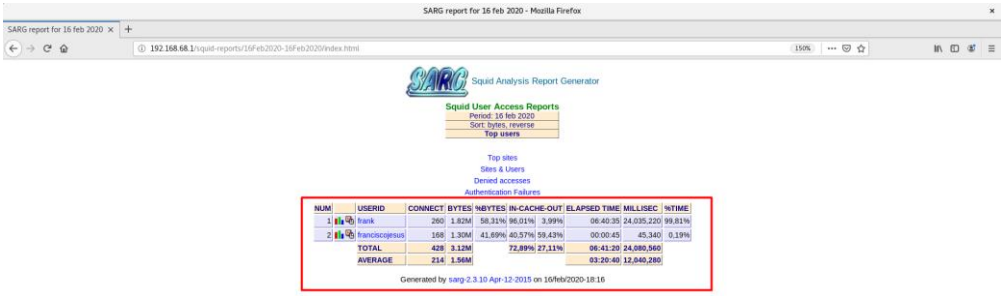
Archivo Editar Ver Buscar Terminal Ayuda

```
root@debian:~# sarg -x
SARG: Init
SARG: Loading configuration from /etc/sarg/sarg.conf
SARG: Chaining IP resolving module "dns"
SARG: Loading exclude host file from: /etc/sarg/exclude_hosts
SARG: Loading exclude file from: /etc/sarg/exclude_users
SARG: Parameters:
SARG:           Hostname or IP address (-a) =
SARG:                    Useragent log (-b) =
SARG:                     Exclude file (-c) = /etc/sarg/exclude_hosts
SARG:                  Date from-until (-d) =
SARG:    Email address to send reports (-e) =
```

Como en la directiva output_dir teníamos configurado el subdirectorio /squid-reports, por defecto accederemos a los informes de Sarg introduciendo la IP/squid-reports.

Podemos ver los informes generados.

Podemos ver las estadísticas de nuestros usuarios.



Podemos ver los sitios más visitados.

Squid Analysis Report Generator

Squid User Access Reports

Period: 16 feb 2020

Sort: bytes, reverse

Top users

Top sites

Sites & Users

Denied accesses

Authentication Failures

NUM		USERID	CONNECT	BYTES	%BYTES	IN-CACHE-OUT	ELAPSED TIME	MILLISEC	%TIME
1		frank	260	1.82M	58,31%	96,01%	3,99%	06:40:35	24,035,220	99,81%
2		franciscojesus	168	1.30M	41,69%	40,57%	59,43%	00:00:45	45,340	0,19%
		TOTAL	**428**	**3.12M**		**72,89%**	**27,11%**	**06:41:20**	**24,080,560**	
		AVERAGE	**214**	**1.56M**				**03:20:40**	**12,040,280**	

Generated by sarg-2.3.10 Apr-12-2015 on 16/feb/2020-18:16

Squid Analysis Report Generator

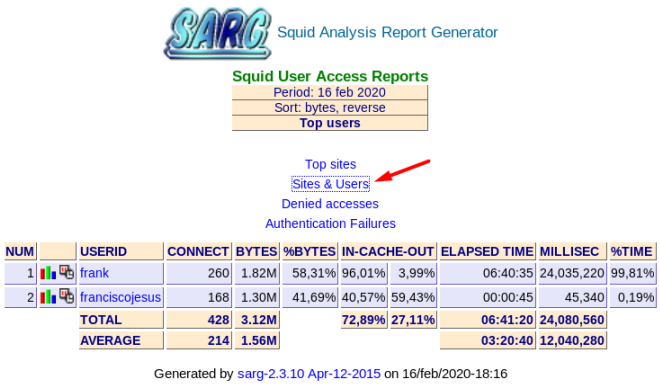
Squid User Access Reports

Period: 16 feb 2020

Top 100 sites

NUM	ACCESSED SITE	CONNECT	BYTES	TIME	USERS
1	www.iesgregorioprieto.com	70	973.26K	0:00:21	2
2	detectportal.firefox.com	70	249.09K	0	1
3	www.google.com:443	68	265.24K	0:00:16	2
4	cdn.syndication.twimg.com:443	44	501.14K	5:11:43	1
5	incoming.telemetry.mozilla.org:443	26	100.97K	0:14:20	1
6	pbs.twimg.com:443	20	276.07K	0:05:44	1
7	debian:3128	20	91.22K	0	2
8	ocsp.digicert.com	16	14.72K	0:00:01	1
9	accounts.google.com:443	12	49.91K	0:00:47	2
10	ocsp.pki.goog	10	8.23K	0:00:03	2
11	safebrowsing.googleapis.com:443	8	249.59K	0:22:47	1
12	csi.gstatic.com	8	1.73K	0:00:03	2
13	syndication.twitter.com:443	6	26.08K	0:06:45	1
14	www.uv.es:443	6	23.97K	0	1
15	www.google-analytics.com:443	6	16.66K	0:05:46	2
16	platform.twitter.com:443	4	142.50K	0:06:34	1

Podemos ver los sitios que más frecuentan los usuarios.



Squid Analysis Report Generator

Squid User Access Reports

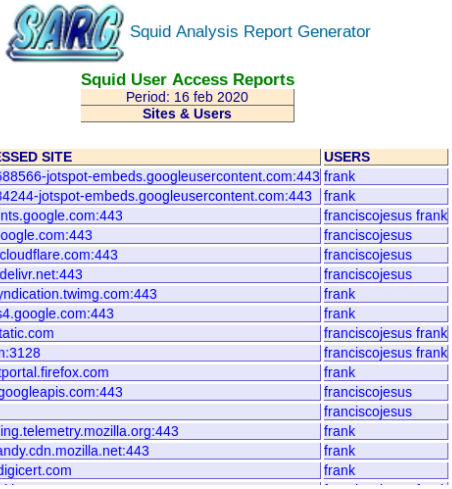
Period: 16 feb 2020
Sort: bytes, reverse
Top users

Top sites
Sites & Users
Denied accesses
Authentication Failures

NUM		USERID	CONNECT	BYTES	%BYTES	IN-CACHE-OUT		ELAPSED TIME	MILLISEC	%TIME
1		frank	260	1.82M	58,31%	96,01%	3,99%	06:40:35	24,035,220	99,81%
2		franciscojesus	168	1.30M	41,69%	40,57%	59,43%	00:00:45	45,340	0,19%
		TOTAL	**428**	**3.12M**		**72,89%**	**27,11%**	**06:41:20**	**24,080,560**	
		AVERAGE	**214**	**1.56M**				**03:20:40**	**12,040,280**	

Generated by sarg-2.3.10 Apr-12-2015 on 16/feb/2020-18:16

Podemos ver los sitios que más visitan los usuarios en el informe generado.



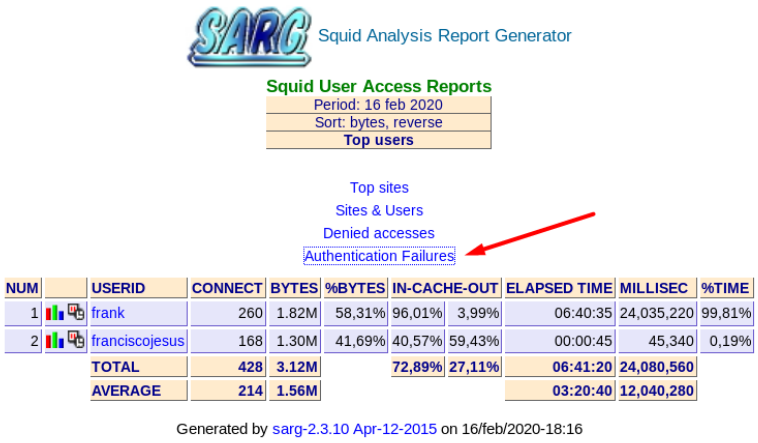
Squid Analysis Report Generator

Squid User Access Reports

Period: 16 feb 2020
Sites & Users

NUM	ACCESSED SITE	USERS
1	1098688566-jotspot-embeds.googleusercontent.com:443	frank
2	540284244-jotspot-embeds.googleusercontent.com:443	frank
3	accounts.google.com:443	franciscojesus frank
4	apis.google.com:443	franciscojesus
5	cdnjs.cloudflare.com:443	franciscojesus
6	cdn.jsdelivr.net:443	franciscojesus
7	cdn.syndication.twimg.com:443	frank
8	clients4.google.com:443	frank
9	csi.gstatic.com	franciscojesus frank
10	debian:3128	franciscojesus frank
11	detectportal.firefox.com	frank
12	fonts.googleapis.com:443	franciscojesus
13	hola	franciscojesus
14	incoming.telemetry.mozilla.org:443	frank
15	normandy.cdn.mozilla.net:443	frank
16	ocsp.digicert.com	frank

Por último, podemos ver las autenticaciones fallidas de cada usuario con su fecha y hora exacta.



Squid Analysis Report Generator

Squid User Access Reports

Period: 16 feb 2020
Sort: bytes, reverse
Top users

Top sites
Sites & Users
Denied accesses
Authentication Failures

NUM		USERID	CONNECT	BYTES	%BYTES	IN-CACHE-OUT		ELAPSED TIME	MILLISEC	%TIME
1		frank	260	1.82M	58,31%	96,01%	3,99%	06:40:35	24,035,220	99,81%
2		franciscojesus	168	1.30M	41,69%	40,57%	59,43%	00:00:45	45,340	0,19%
		TOTAL	**428**	**3.12M**		**72,89%**	**27,11%**	**06:41:20**	**24,080,560**	
		AVERAGE	**214**	**1.56M**				**03:20:40**	**12,040,280**	

Generated by sarg-2.3.10 Apr-12-2015 on 16/feb/2020-18:16

Squid Analysis Report Generator

Squid User Access Reports

Period: 16 feb 2020

Authentication Failures

USERID	IP/NAME	DATE/TIME	ACCESSED SITE
192.168.68.10	192.168.68.10	16/02/20-15:50:06	accounts.google.com:443
		16/02/20-15:50:06	accounts.google.com:443
		16/02/20-15:50:07	accounts.google.com:443
		16/02/20-15:50:07	accounts.google.com:443
		16/02/20-15:50:10	accounts.google.com:443
		16/02/20-15:50:10	accounts.google.com:443
		16/02/20-15:50:19	accounts.google.com:443
		16/02/20-15:50:19	accounts.google.com:443
		16/02/20-15:52:40	accounts.google.com:443
		16/02/20-15:52:40	accounts.google.com:443
			282 more authentication failures not shown here...
192.168.68.20	192.168.68.20	16/02/20-16:03:07	apis.google.com:443
		16/02/20-16:03:07	apis.google.com:443
		16/02/20-16:03:04	content-signature-2.cdn.mozilla.net:443
		16/02/20-16:03:04	content-signature-2.cdn.mozilla.net:443
		16/02/20-16:03:04	content-signature-2.cdn.mozilla.net:443
		16/02/20-16:03:04	content-signature-2.cdn.mozilla.net:443
		16/02/20-16:04:05	content-signature-2.cdn.mozilla.net:443
		16/02/20-16:04:05	content-signature-2.cdn.mozilla.net:443
		16/02/20-16:10:39	content-signature-2.cdn.mozilla.net:443
		16/02/20-16:10:39	content-signature-2.cdn.mozilla.net:443
			168 more authentication failures not shown here...

Generated by sarg-2.3.10 Apr-12-2015 on 16/feb/2020-18:16

Si pulsamos en un usuario podemos ver más información acerca de lo que este realizo.

Squid Analysis Report Generator

Squid User Access Reports

Period: 16 feb 2020

Sort: bytes, reverse

Top users

Top sites

Sites & Users

Denied accesses

Authentication Failures

NUM		USERID	CONNECT	BYTES	%BYTES	IN-CACHE-OUT		ELAPSED TIME	MILLISEC	%TIME
1		frank	260	1.82M	58,31%	96,01%	3,99%	06:40:35	24,035,220	99,81%
2		franciscojesus	168	1.30M	41,69%	40,57%	59,43%	00:00:45	45,340	0,19%
		TOTAL	**428**	**3.12M**		**72,89%**	**27,11%**	**06:41:20**	**24,080,560**	
		AVERAGE	**214**	**1.56M**				**03:20:40**	**12,040,280**	

Generated by sarg-2.3.10 Apr-12-2015 on 16/feb/2020-18:16

Squid Analysis Report Generator

Squid User Access Reports

Period: 16 feb 2020

User: franciscojesus

Sort: bytes, reverse

User report

	ACCESSED SITE	CONNECT	BYTES	%BYTES	IN-CACHE-OUT		ELAPSED TIME	MILLISEC	%TIME	
	www.iesgregorioprieto.com	50	864.46K	66,30%	10,75%	89,25%	00:00:18	18,436	40,66%	
	www.google.com:443	64	257.11K	19,72%	100,00%	0,00%	00:00:11	11,942	26,34%	
	debian:3128	12	54.86K	4,21%	100,00%	0,00%	00:00:00	0	0,00%	DENIED
	accounts.google.com:443	10	40.26K	3,09%	100,00%	0,00%	00:00:00	4	0,01%	DENIED
	www.uv.es:443	6	23.97K	1,84%	100,00%	0,00%	00:00:00	0	0,00%	DENIED
	uv.es	4	17.74K	1,36%	100,00%	0,00%	00:00:00	0	0,00%	DENIED
	apis.google.com:443	2	9.24K	0,71%	100,00%	0,00%	00:00:11	11,918	26,29%	
	update.googleapis.com:443	2	8.06K	0,62%	100,00%	0,00%	00:00:00	0	0,00%	DENIED
	fonts.googleapis.com:443	2	8.05K	0,62%	100,00%	0,00%	00:00:00	0	0,00%	DENIED
	cdnjs.cloudflare.com:443	2	8.05K	0,62%	100,00%	0,00%	00:00:00	0	0,00%	DENIED
	cdn.jsdelivr.net:443	2	8.03K	0,62%	100,00%	0,00%	00:00:00	0	0,00%	DENIED
	ocsp.pki.goog	4	3.29K	0,25%	0,00%	100,00%	00:00:02	2,074	4,57%	
	hola	2	646	0,05%	100,00%	0,00%	00:00:00	0	0,00%	DENIED
	csi.gstatic.com	4	0	0,00%	0,00%	0,00%	00:00:00	668	1,47%	
	www.google-analytics.com:443	2	0	0,00%	0,00%	0,00%	00:00:00	298	0,66%	
	TOTAL	**168**	**1.30M**	**41,69%**	**40,57%**	**59,43%**	**00:00:45**	**45,340**	**0,19%**	
	AVERAGE	**0**	**1.56M**				**03:20:40**	**12,040,280**	**50,00%**	

Generated by sarg-2.3.10 Apr-12-2015 on 16/feb/2020-18:16

Para finalizar, podemos ver gráficas de los usuarios.

Squid Analysis Report Generator

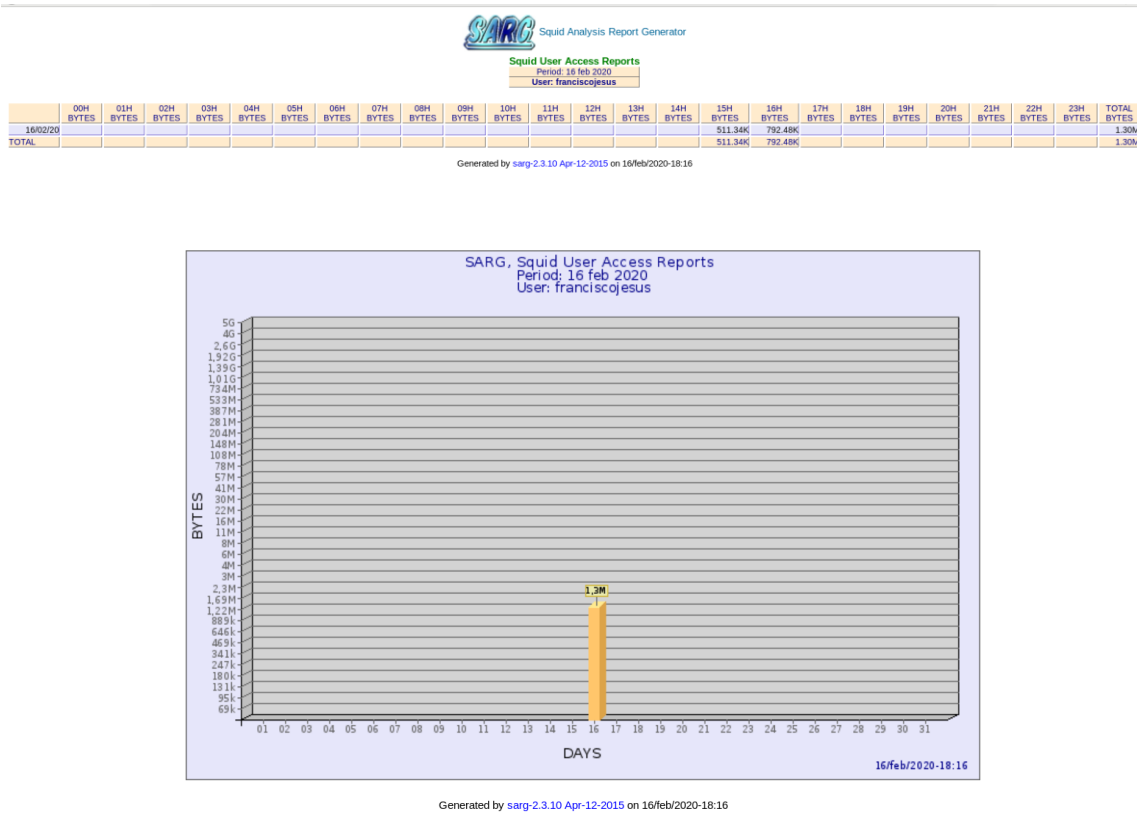
Squid User Access Reports

Period: 16 feb 2020
Sort: bytes, reverse
Top users

Top sites
Sites & Users
Denied accesses
Authentication Failures

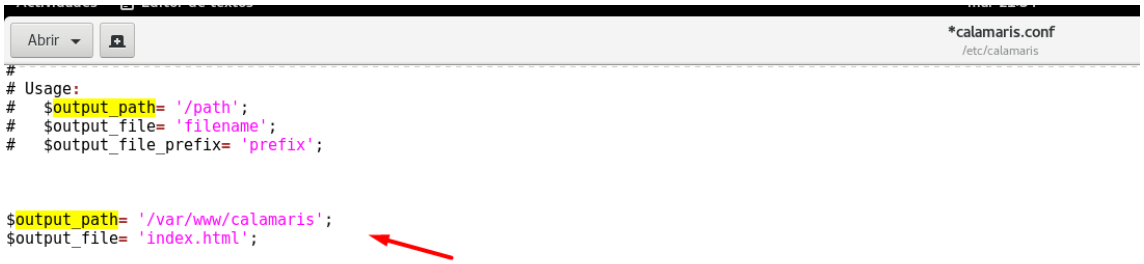
NUM		USERID	CONNECT	BYTES	%BYTES	IN-CACHE-OUT		ELAPSED TIME	MILLISEC	%TIME
1		frank	260	1.82M	58,31%	96,01%	3,99%	06:40:35	24,035,220	99,81%
2		franciscojesus	168	1.30M	41,69%	40,57%	59,43%	00:00:45	45,340	0,19%
		TOTAL	**428**	**3.12M**		**72,89%**	**27,11%**	**06:41:20**	**24,080,560**	
		AVERAGE	**214**	**1.56M**				**03:20:40**	**12,040,280**	

Generated by sarg-2.3.10 Apr-12-2015 on 16/feb/2020-18:16

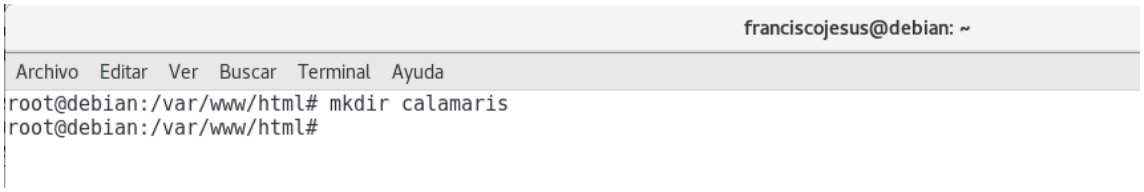


Calamaris es un script de Perl que se usa para generar informes de actividad de caché en formato ASCII o HTML. Funciona con archivos de registro de acceso de Squid nativos.

Vamos a su archivo de configuración y lo configuramos con el archivo log de squid.



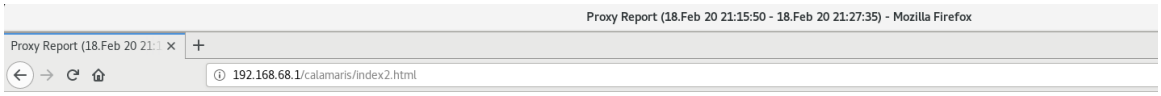
Creamos la carpeta de calamaris.



Usamos el siguiente comando para generar el informe en Calamaris (este podría ser automatizado con crontab).

```
cat /var/log/squid/access.log | calamaris -a -F html >
/var/www/html/calamaris/index.html
```

Ya podemos acceder vía web.



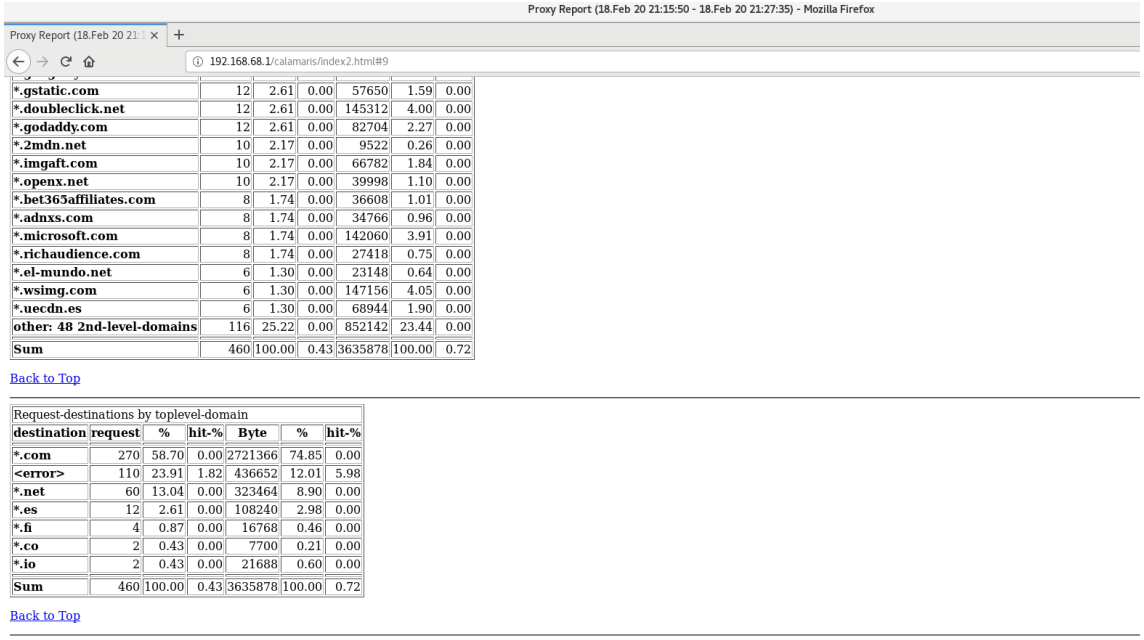
Proxy Report

Report period: **18.Feb 20 21:15:50 - 18.Feb 20 21:27:35**
Generated at: 18.Feb 20 21:39:27

Table of Content / Overview			
Summary	-	-	-
Incoming requests by method	most requested method	CONNECT	398 Requests
Incoming UDP-requests by status	-	-	no requests found
Incoming TCP-requests by status	most incoming request by status to	MISS	350 Requests
Outgoing requests by status	most outgoing request to	DIRECT Fetch from Source	350 Requests
Outgoing requests by destination	most requested destination	DIRECT	350 Requests
Request-destinations by 2nd-level-domain	most requested 2nd-level-domain	<error>	110 Requests
Request-destinations by toplevel-domain	most requested toplevel-domain	*.com	270 Requests
TCP-Request-protocol	most requested protocol	<secure>	310 Requests
Requested content-type	most requested content-type	<secure>	310 Requests
Requested extensions	most requested extension	<secure>	310 Requests
Incoming UDP-requests by host	-	-	no requests found
Incoming TCP-requests by host	most active host	192.168.68.10	460 Requests
Size Distribution Diagram	most requested object_size	1000-9999	352 Requests
Performance in 1 hour steps	most active day	18.Feb 20 21:00	460 Requests
UDP-Request duration distribution in msec	-	-	no requests found
TCP-Request duration distribution in msec	most frequent response time	<= 0.1	94 Requests
UDP Response code distribution	-	-	no requests found
TCP Response code distribution	most frequent response code	200	350 Requests

Summary		
Calamaris statistics		
lines parsed:	lines	460
invalid lines:	lines	0
parse time:	sec	0

Podemos ver los dominios de primer nivel y subdominios más buscados.



***.gstatic.com**	12	2.61	0.00	57650	1.59	0.00
***.doubleclick.net**	12	2.61	0.00	145312	4.00	0.00
***.godaddy.com**	12	2.61	0.00	82704	2.27	0.00
***.2mdn.net**	10	2.17	0.00	9522	0.26	0.00
***.imgaft.com**	10	2.17	0.00	66782	1.84	0.00
***.openx.net**	10	2.17	0.00	39998	1.10	0.00
***.bet365affiliates.com**	8	1.74	0.00	36608	1.01	0.00
***.adnxs.com**	8	1.74	0.00	34766	0.96	0.00
***.microsoft.com**	8	1.74	0.00	142060	3.91	0.00
***.richaudience.com**	8	1.74	0.00	27418	0.75	0.00
***.el-mundo.net**	6	1.30	0.00	23148	0.64	0.00
***.wsimg.com**	6	1.30	0.00	147156	4.05	0.00
***.uecdn.es**	6	1.30	0.00	68944	1.90	0.00
other: 48 2nd-level-domains	116	25.22	0.00	852142	23.44	0.00
Sum	460	100.00	0.43	3635878	100.00	0.72

Back to Top

Request-destinations by toplevel-domain

destination	request	%	hit-%	Byte	%	hit-%
***.com**	270	58.70	0.00	2721366	74.85	0.00
<error>	110	23.91	1.82	436652	12.01	5.98
***.net**	60	13.04	0.00	323464	8.90	0.00
***.es**	12	2.61	0.00	108240	2.98	0.00
***.fi**	4	0.87	0.00	16768	0.46	0.00
***.co**	2	0.43	0.00	7700	0.21	0.00
***.io**	2	0.43	0.00	21688	0.60	0.00
Sum	460	100.00	0.43	3635878	100.00	0.72

Back to Top