Trabalho Laboratorial 1

Planeamento e Gestão de Redes

Sala I321 Bancada 1

Diogo Remião & Miguel Pinheiro

FACULDADE DE ENGENHARIA

UNIVERSIDADE DO PORTO

Faculdade de Engenharia da Universidade do Porto
TEC

Contents

1 Setup

Neste capítulo vamos abordar o setup de teste que foi utilizado neste trabalho e como foram configuradas as diferentes ferramentas, nomeadamente o Apache e o Squid Cache.

1.1 Servers HTTP

1.1.1 Apache

Neste trabalho é necessário configurar dois websites dinâmicos que vão depois ser acedidos. E necessário que se gerem logs de acesso que irão futuramente ser analisados. Neste sentido, foi utilizada a ferramenta sugerida pelo docente, o Apache.

O Apache é um servidor httpd open-source que permite dar deploy de uma forma simples a sites web a partir do computador do utilizador. No nosso caso, configuramos dois websites, intuitivamente apelidados de site1 e site2, que foram alojados no tux 13 (172.16.1.13). O primeiro site está designado à porta 80, porta *default* para acessos HTTP, e o segundo na porta 81.

Desta forma, para aceder aos websites basta colocar no browser de um computador na mesma rede (irá ser abordado em 1.2) o IP 172.16.1.13 e 172.16.1.13:81 respetivamente. Note-se que se coloca ":81" para forçar o browser a aceder por essa porta. Caso contrário, iria optar pela porta predefinida para HTTP.

1.1.2 Dynamic Update

Hoje em dia é comum implementações de serviços de cache que guardam informação num servidor mais próximo do utilizador, muitas vezes no próprio computador. Este serviço serve para diminuir a carga no *host* do website. Apesar de ser útil em contexto real, neste trabalho não o é pois impede o registo de todos os acessos aos websites. Desta forma, os sites são carregados com conteúdo dinâmico que obriga sempre a aceder ao site através do seu *host*. Foi utilizado o code snippet fornecido pelo docente programado em PHP, pelo que foi necessário ativar esta ferramenta no Apache.

1.1.3 Logging

Tal como indicado previamente, o Apache permite o registo em *logs* dos acessos feitos aos websites. Estes logs são cruciais pois vão ser analisados futuramente em 2. Foram criados dois diretórios diferentes para cada site, cada um contendo o respetivo *access.log* e *error.log*.

1.2 Proxy HTTP

1.2.1 Squid Cache

O Squid Cache é um proxy (com funcionalidades de cache) open-source que mais um vez permite alojar um proxy num *host*. Os acessos proxy são geralmente feitos através da porta 3128. Neste caso, o IP do computador a alojar o proxy é 192.168.109.11. Por isso, para aceder ao proxy, basta escolhermos como IP do proxy 192.168.109.11:3128.

Uma outra funcionalidade do Squid neste trabalho prende-se a estrutura da rede em que este trabalho está a ser desenvolvido. O servidor tux onde estão a ser alojados os websites é um subdomain de servidor de bancada onde está a ser alojado o proxy. Desta forma, mesmo dentro do VPN institucional, o nosso computador pessoal não consegue aceder diretamente aos sites. Para isso é preciso usar o Squid que funciona como um túnel. Desta forma, temos acesso direto aos websites a partir de um computador fora da rede dos tux.

1.2.2 Logging

Tal como o Apache, o Squid também mantém um registo de *logs* com todos os acessos feitos pelo squid, incluíndo o IP de origem e de destino. Estes logs irão ser futuramente analisados também.

2 Test Environment

Nesta secção vamos explicar como se procedeu para obter os resultados neste trabalho.

2.1 Setup

Tal como indicado previamente, os dois servidores web estão alojados no tux 172.16.1.13, nas portas 80 e 81.

O proxy Squid tem por predefinição a porta 80 aberta mas a porta 81 teve que ser manualmente aberta no ficheiro de configuração do Squid, de forma a permitir o acesso ao segundo site.

Os servidores web foram acedidos periodicamente por um outro tux dentro da mesma rede, o 172.16.1.14.

2.2 Logging e Cronjobs

Os logs de acesso do Apache irão mostrar diferentes tipos de acessos aos dois websites. Em primeiro lugar, foram feitos acessos de um computador externo ligado pela rede VPN. Estes acessos foram feitos com o intuito de testar o funcionamento das diferentes ferramentas.

Para criar acessos periódicos que criem logs com mais informação útil, foram configurados acessos periódicos do tux 172.16.1.14. Este acessos ocorriam inicialmente uma vez por hora durante um dia e 4 vezes por hora no segundo dia. Os acessos eram feitos quer diretamente, quer através do proxy, de forma a gerar mais informação. Para este efeito forma adicionados Cronjobs à crontab do tux, que são nada mais nada menos que comandos executados pelo computador de forma autónoma nos *timestamps* desejados

A razão pela qual não foram gerados mais logs foi o facto de no inicio da semana estes Cronjobs terem sido configurados para correr, e devido a problemas técnicos, houve um wipeout de vários serviços, nomeadamente do proxy. Tivemos então que recomeçar o teste mas com menos tempo para gerar os logs devido à data limite de entrega do trabalho. De qualquer modo, a informação gerada foi suficiente para obter resultados úteis e proceder à sua análise.

3 Análise dos Logs

Nesta secção iremos analisar diferentes ferramentas de análise de *logs* e compará-las em termos de funcionalidade.

3.1 Webalizer

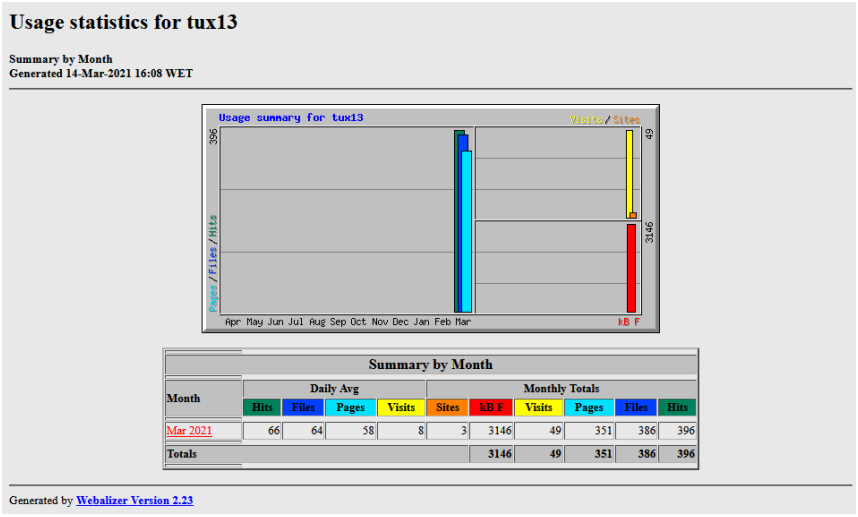
Webalizer é uma ferramenta open-source de análise de logs. Possui um ficheiro de configuração que o permite suportar várias *features*. A sua instalação é bastante intuitiva e para este trabalho, o ficheiro de configuração não foi alterado com exceção dos necessário para o programa reconhecer os logs necessários.

Depois de dar deploy ao Webalizer, podemos fazer aceder ao site criado pelo browser pessoal através do proxy. De notar que Webalizer é um subdomain do site por isso o *host* é o mesmo, neste caso o tux.

Ao abrir o Webalizer, este apresenta um resumo anual do tráfico gerado nesse site (Fig 3.1). Neste caso, como todo o tráfico gerado foi no mês de Março, é apenas esse que aparece. Apresenta uma média diário e um total mensal para cada mês. Como se pode ver, são distinguidos vários tipos de acessos ao website:

- Hits: Representa todos os *requests* que foram enviados ao website, daí que o seu número seja o maior.
- Files: Representa todos os *requests* em que o servidor envia alguma coisa ao client. Em suma, os files são as respostas aos hits por isso o seu número deve ser próximo.
- Pages: Representa todos os *requests* de páginas, neste caso php.
- Sites: Representa as diferentes origens de *requests* ao servidor, ou seja, os diferentes utilizadores que acederam.
- Visits: A partir do momento em que um *request* é feito ao servidor, uma visita é iniciada. Todos os pedidos feitos ao site sem que o timeout expire representam apenas uma visita.
- Kbytes: Representa o total de kBytes enviados pelo servidor.

As diferenças podem ser melhor analisadas e justificadas com a análise dos dados mais detalhada fornecida pelo Webalizer. Ao clicar no mês de Março, o Webalizer fornece um vista mais detalhada desse mês.



Usage statistics for tux13

Summary by Month
Generated 14-Mar-2021 16:08 WET

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	kB F	Visits	Pages	Files	Hits
Mar 2021	66	64	58	8	3	3146	49	351	386	396
Totals						3146	49	351	386	396

Generated by Webalizer Version 2.23

Figure 3.1: Main Page do Webalizer (site1)

Monthly Statistics for March 2021		
Total Hits		396
Total Files		386
Total Pages		351
Total Visits		49
Total kB Files		3146
Total Unique Sites		3
Total Unique URLs		35
Total Unique Referrers		4
Total Unique User Agents		2
	Avg	Max
Hits per Hour	2	49
Hits per Day	66	181
Files per Day	64	178
Pages per Day	58	178
Sites per Day	0	3
Visits per Day	8	34
kB Files per Day	524	1714
Hits by Response Code		
Code 200 - OK	97.47%	386
Code 404 - Not Found	2.53%	10

Figure 3.2: Análise Mensal do Webalizer (site1)

Na figura 3.2 podemos mais uma vez ver a informação apresentada na página principal, assim como uma média horária e diária dos diferentes tipos de acesso. Algumas conclusões podem ser retiradas com estes dados:

- A diferença entre os "Hits" e os "Files" é de 10 unidades, que foram erros 404 na durante o processo de configuração dos servidores.
- O valores máximos são muitos superiores aos da média. Isto deve-se ao facto de durante a configuração dos servidores se terem feito muitos testes usando o firefox que recarregava as páginas a cada 5 segundos. Assim, rapidamente se obtém valores de acesso muito elevados.
- O número médio de acessos por hora é de 2. Isto vai ao encontro dos acessos periódicos definidos no tux, 1 por hora no primeiro dia e 4 por hora no segundo, que em média é aproximadamente 2. Tenha-se em atenção acessos não constantes que influenciaram este valor.

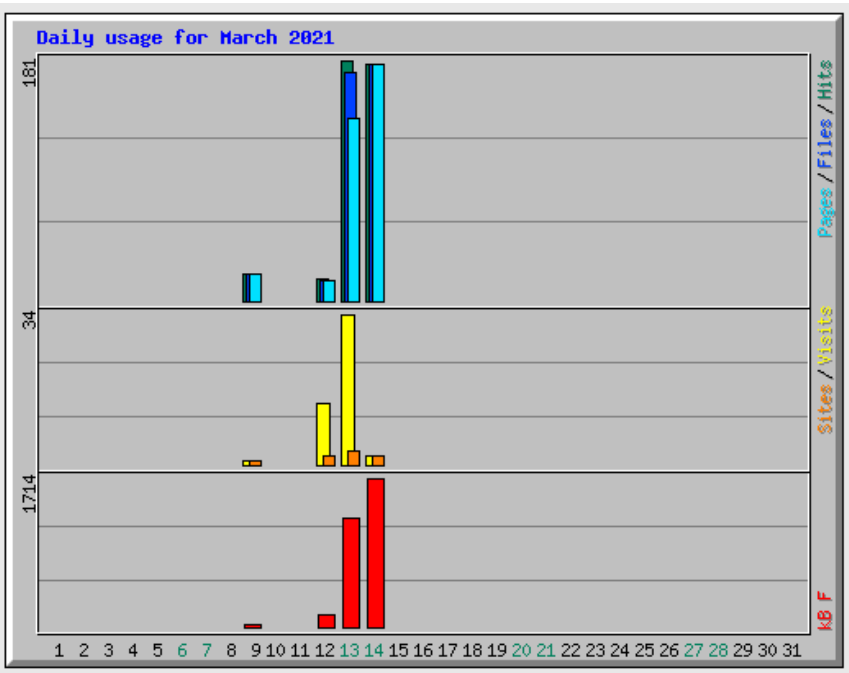
O Webalizer também apresenta uma análise diária dos acessos, com um gráfico onde podemos observar os dias com mais ou menos movimento (Fig 3.3). Os dias 13 e 14 foram os dias em que o tux esteve a fazer os acessos periódicos, sendo que podemos observar que no dia 14 houve mais acessos, tal como esperado. Se esta informação fosse retirada no fim do dia, seria de esperar um número de acessos quatro vezes superior. Os restantes dias correspondem aos acessos feitos durante as configurações.

Uma análise semelhante é feita mas horária, onde os resultados são apresentados em função das horas. Este gráfico é particularmente interessante para avaliar as horas de maior intensidade de acessos aos websites (Fig 3.4). Conseguimos ver a que horas é que estivemos a trabalhar para configurar o proxy :)

Por fim, na figura 3.5 podemos observar quais foram os diretórios do site mais acedidos. Dado que apenas configuramos acessos ao main, este representa 97%, com os restantes a serem as ferramentas de análise que foram acedidas apenas no fim.

Podemos também ver os IPs dos *hosts* que acederam ao website e as respetivas percentagens. Neste caso temos 3:

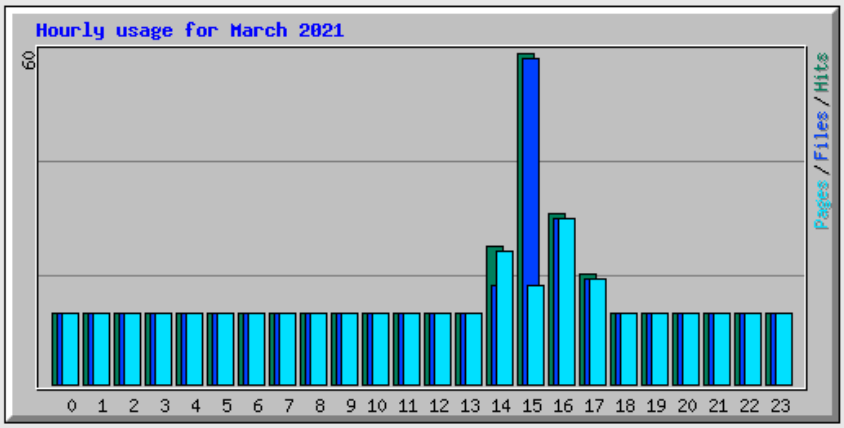
- O acesso pelo proxy, usado quer pelos tux que por nós no nosso computador de casa
- O acesso direto dentro da rede por outro tux
- O acesso *localhost*, feito a partir do tux13 que estava a dar host



Daily Statistics for March 2021

Day	Hits		Files		Pages		Visits		Sites		kB F	
9	20	5.05%	20	5.18%	20	5.70%	1	2.04%	1	33.33%	31	0.99%
10	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
11	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
12	17	4.29%	16	4.15%	16	4.56%	14	28.57%	2	66.67%	145	4.61%
13	181	45.71%	172	44.56%	137	39.03%	34	69.39%	3	100.00%	1256	39.92%
14	178	44.95%	178	46.11%	178	50.71%	2	4.08%	2	66.67%	1714	54.48%

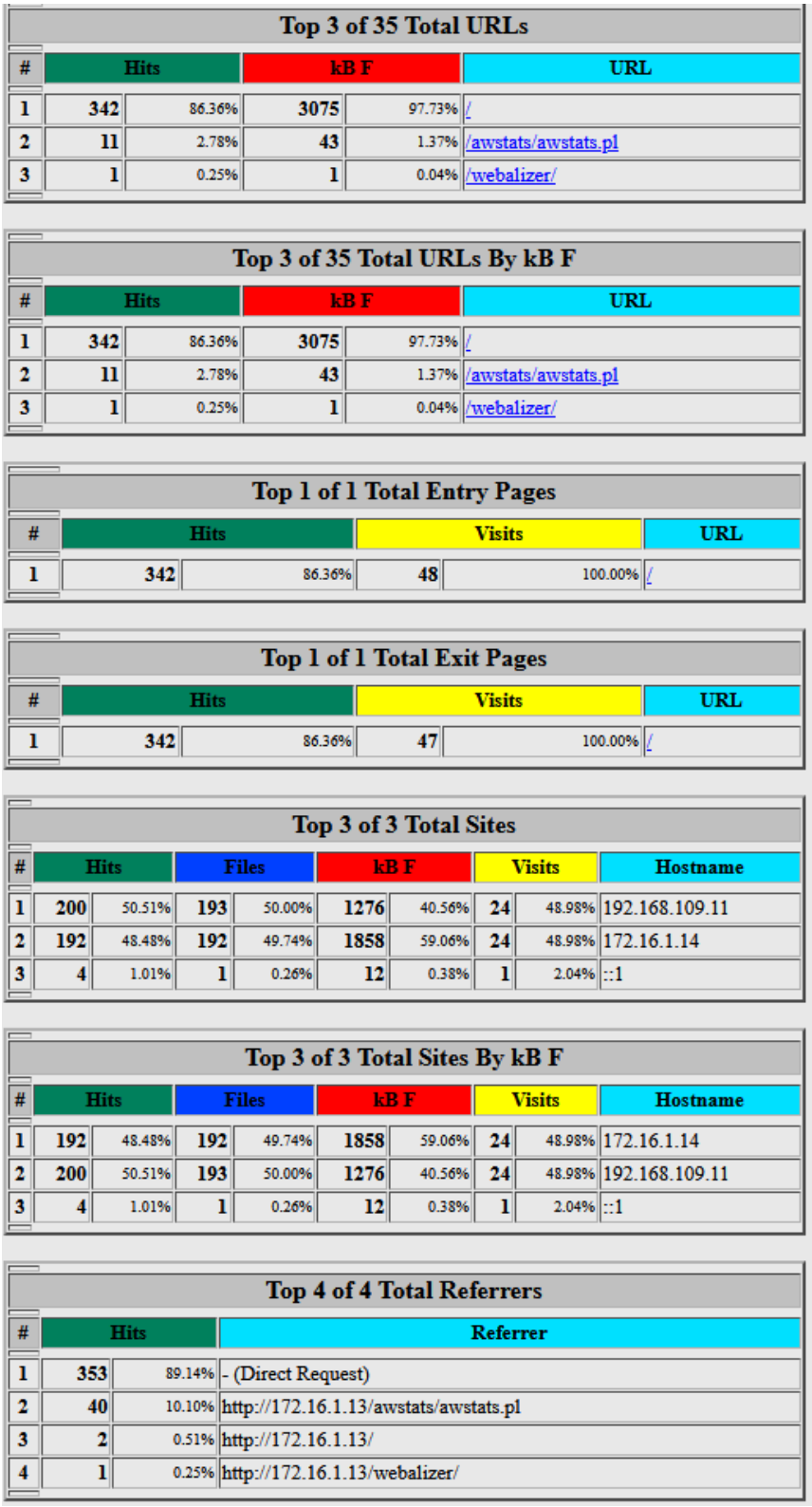
Figure 3.3: Análise diária do Webalizer (site1)



Hourly Statistics for March 2021

Hour	Hits			Files			Pages			kB F		
	Avg	Total		Avg	Total		Avg	Total		Avg	Total	
0	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	126	3.99%
1	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
2	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
3	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
4	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
5	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
6	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
7	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	126	3.99%
8	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
9	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
10	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
11	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
12	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
13	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
14	4	25	6.31%	3	18	4.66%	4	24	6.84%	24	145	4.62%
15	10	60	15.15%	9	59	15.28%	3	18	5.13%	33	196	6.24%
16	5	31	7.83%	5	30	7.77%	5	30	8.55%	27	161	5.11%
17	3	20	5.05%	3	19	4.92%	3	19	5.41%	23	136	4.33%
18	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	127	4.02%
19	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	127	4.02%
20	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	127	4.02%
21	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	127	4.02%
22	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%
23	2	13	3.28%	2	13	3.37%	2	13	3.70%	21	125	3.97%

Figure 3.4: Análise horária do Webalizer (site1)



Top 3 of 35 Total URLs

#	Hits		kB F		URL
1	342	86.36%	3075	97.73%	/
2	11	2.78%	43	1.37%	/awstats/awstats.pl
3	1	0.25%	1	0.04%	/webalizer/

Top 3 of 35 Total URLs By kB F

#	Hits		kB F		URL
1	342	86.36%	3075	97.73%	/
2	11	2.78%	43	1.37%	/awstats/awstats.pl
3	1	0.25%	1	0.04%	/webalizer/

Top 1 of 1 Total Entry Pages

#	Hits		Visits		URL
1	342	86.36%	48	100.00%	/

Top 1 of 1 Total Exit Pages

#	Hits		Visits		URL
1	342	86.36%	47	100.00%	/

Top 3 of 3 Total Sites

#	Hits		Files		kB F		Visits		Hostname
1	200	50.51%	193	50.00%	1276	40.56%	24	48.98%	192.168.109.11
2	192	48.48%	192	49.74%	1858	59.06%	24	48.98%	172.16.1.14
3	4	1.01%	1	0.26%	12	0.38%	1	2.04%	::1

Top 3 of 3 Total Sites By kB F

#	Hits		Files		kB F		Visits		Hostname
1	192	48.48%	192	49.74%	1858	59.06%	24	48.98%	172.16.1.14
2	200	50.51%	193	50.00%	1276	40.56%	24	48.98%	192.168.109.11
3	4	1.01%	1	0.26%	12	0.38%	1	2.04%	::1

Top 4 of 4 Total Referrers

#	Hits		Referrer
1	353	89.14%	- (Direct Request)
2	40	10.10%	http://172.16.1.13/awstats/awstats.pl
3	2	0.51%	http://172.16.1.13/
4	1	0.25%	http://172.16.1.13/webalizer/

Figure 3.5: Outras informações do Webalizer (site1)

Figure 3.6: Topo da página do AWStats (site2)

3.2 AWStats

O Advanced Web Statistics, mais conhecido como o AWStats, é uma ferramenta de análise de logs capaz de gerar gráficos que apresentação de uma forma visual e intuitiva as estatísticas de um website.

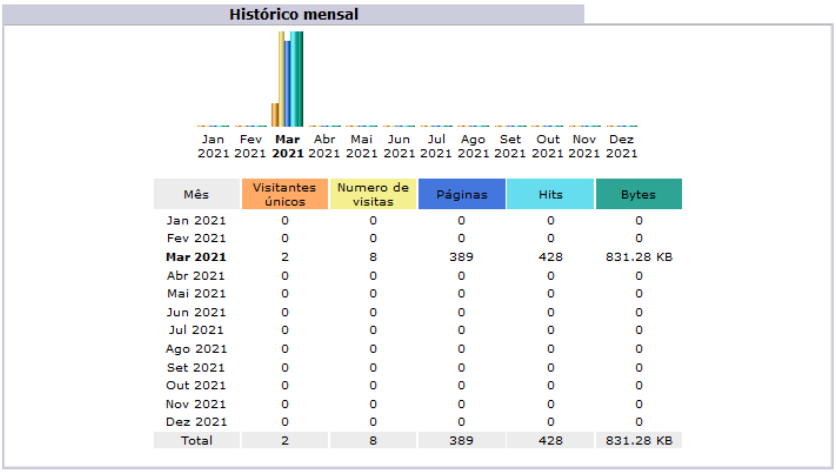
Após a instalação desta ferramenta, foi necessário fazer algumas alterações a nível de configuração antes de poder dar deploy do AWStats. Foi necessário especificar no ficheiro de configuração a localização dos logs dos nossos websites e os seus domínios, no nosso caso através do seu IP 172.16.1.13 e 172.16.1.13:81. Ao abrir o AWStats , é-nos apresentado um resumo do tráfego gerado no presente mês , neste caso de Março de 2021 (Fig 3.6).

São apresentadas várias estatísticas como:

- Visitantes únicos: Pessoa ou computador que fez pelo menos 1 "Hit" em 1 página do website durante o período considerado. São distinguidos através do seu IP.
- Número de visitas: Número de visitas feitas por todos os visitantes. Vários acessos pelo mesmo IP contam como uma visita dependendo do timeout definido.
- Páginas: Representa todos os *requests* de páginas, neste caso php.
- Hits: Representa todos os *requests* que foram enviados ao website.
- Bytes: Representa o total de Bytes enviados pelo servidor.

O gráfico mensal possui informação muito semelhante à página inicial (Fig 3.7). Como o nosso trabalho foi realizado durante apenas o mês de março , apenas este possui estatística.

Analisando as estatísticas diárias (Fig 3.8) , conseguimos distinguir um maior número de acessos nos dias 13 e 14 de Março , que representam os dias em que o tux esteve a realizar acessos periódicos . Verifica-se uma média diária muito baixa fruto do acesso periódico dos tux's ter sido realizado apenas durante 2 dias e nos

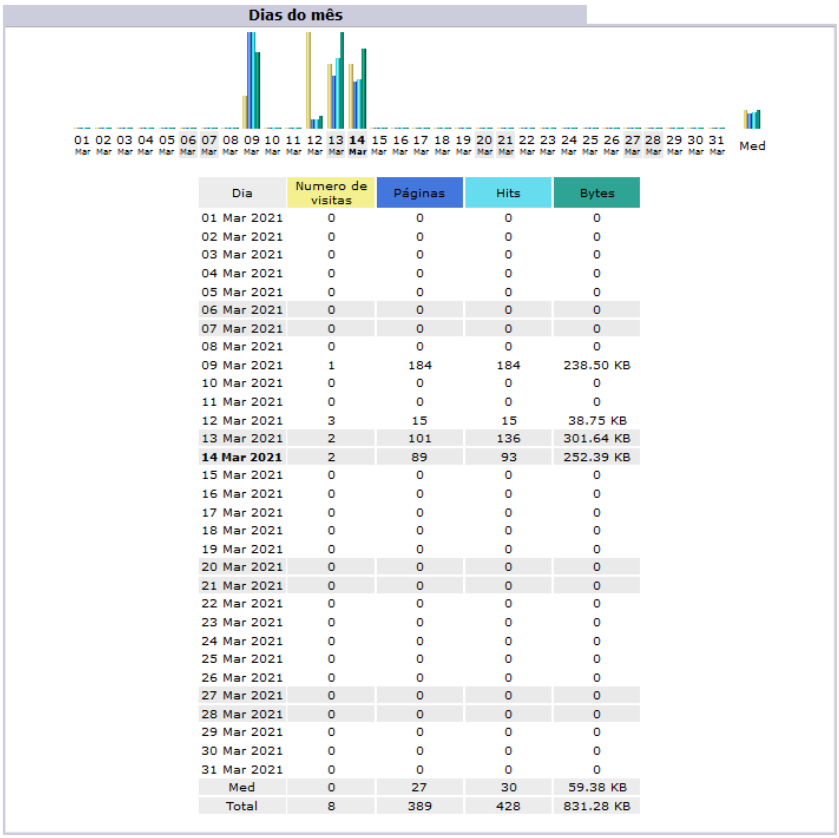


Histórico mensal

Jan 2021 Fev 2021 **Mar 2021** Abr 2021 Mai 2021 Jun 2021 Jul 2021 Ago 2021 Set 2021 Out 2021 Nov 2021 Dez 2021

Mês	Visitantes únicos	Numero de visitas	Páginas	Hits	Bytes
Jan 2021	0	0	0	0	0
Fev 2021	0	0	0	0	0
Mar 2021	2	8	389	428	831.28 KB
Abr 2021	0	0	0	0	0
Mai 2021	0	0	0	0	0
Jun 2021	0	0	0	0	0
Jul 2021	0	0	0	0	0
Ago 2021	0	0	0	0	0
Set 2021	0	0	0	0	0
Out 2021	0	0	0	0	0
Nov 2021	0	0	0	0	0
Dez 2021	0	0	0	0	0
Total	2	8	389	428	831.28 KB

Figure 3.7: Análise Mensal do AWStats (site2)

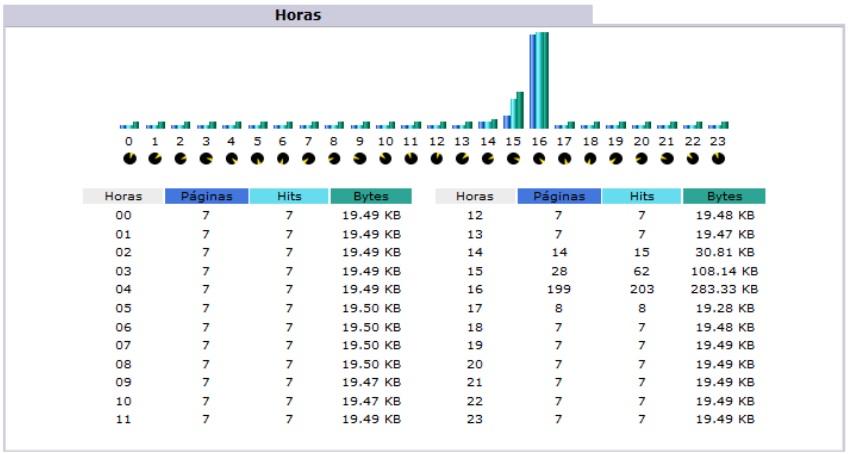


Dias do mês

01 Mar 02 Mar 03 Mar 04 Mar 05 Mar 06 Mar 07 Mar 08 Mar 09 Mar 10 Mar 11 Mar 12 Mar 13 Mar **14 Mar** 15 Mar 16 Mar 17 Mar 18 Mar 19 Mar 20 Mar 21 Mar 22 Mar 23 Mar 24 Mar 25 Mar 26 Mar 27 Mar 28 Mar 29 Mar 30 Mar 31 Mar Med

Dia	Numero de visitas	Páginas	Hits	Bytes
01 Mar 2021	0	0	0	0
02 Mar 2021	0	0	0	0
03 Mar 2021	0	0	0	0
04 Mar 2021	0	0	0	0
05 Mar 2021	0	0	0	0
06 Mar 2021	0	0	0	0
07 Mar 2021	0	0	0	0
08 Mar 2021	0	0	0	0
09 Mar 2021	1	184	184	238.50 KB
10 Mar 2021	0	0	0	0
11 Mar 2021	0	0	0	0
12 Mar 2021	3	15	15	38.75 KB
13 Mar 2021	2	101	136	301.64 KB
14 Mar 2021	2	89	93	252.39 KB
15 Mar 2021	0	0	0	0
16 Mar 2021	0	0	0	0
17 Mar 2021	0	0	0	0
18 Mar 2021	0	0	0	0
19 Mar 2021	0	0	0	0
20 Mar 2021	0	0	0	0
21 Mar 2021	0	0	0	0
22 Mar 2021	0	0	0	0
23 Mar 2021	0	0	0	0
24 Mar 2021	0	0	0	0
25 Mar 2021	0	0	0	0
26 Mar 2021	0	0	0	0
27 Mar 2021	0	0	0	0
28 Mar 2021	0	0	0	0
29 Mar 2021	0	0	0	0
30 Mar 2021	0	0	0	0
31 Mar 2021	0	0	0	0
Med	0	27	30	59.38 KB
Total	8	389	428	831.28 KB

Figure 3.8: Análise Diária do AWStats (site2)



Horas	Páginas	Hits	Bytes	Horas	Páginas	Hits	Bytes
00	7	7	19.49 KB	12	7	7	19.48 KB
01	7	7	19.49 KB	13	7	7	19.47 KB
02	7	7	19.49 KB	14	14	15	30.81 KB
03	7	7	19.49 KB	15	28	62	108.14 KB
04	7	7	19.49 KB	16	199	203	283.33 KB
05	7	7	19.50 KB	17	8	8	19.28 KB
06	7	7	19.50 KB	18	7	7	19.48 KB
07	7	7	19.50 KB	19	7	7	19.49 KB
08	7	7	19.50 KB	20	7	7	19.49 KB
09	7	7	19.47 KB	21	7	7	19.49 KB
10	7	7	19.47 KB	22	7	7	19.49 KB
11	7	7	19.49 KB	23	7	7	19.49 KB

Figure 3.9: Análise Horária do AWStats (site2)

restantes dias apenas eram realizados alguns acessos durante a configuração dos websites.

É de notar que AwStats também apresenta uma análise horária semelhante à do Webalizer (Fig 3.9).

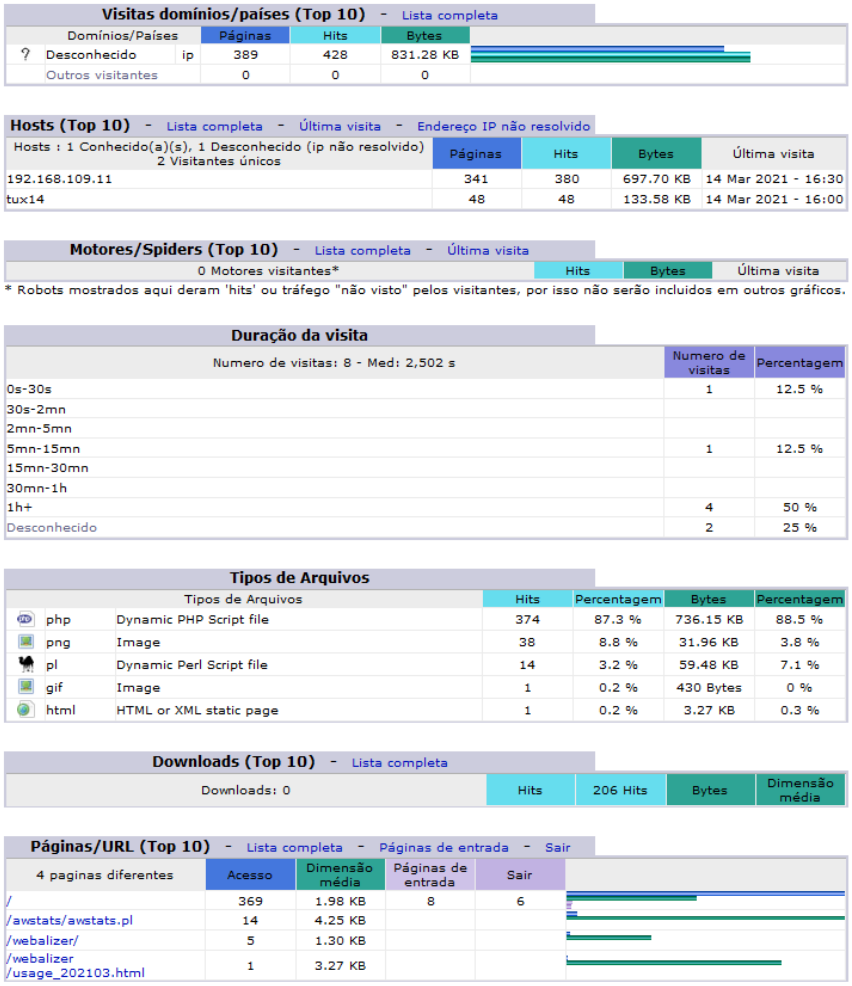
Por fim, na figura 3.10 é apresentada informação variada. Primeiramente um TOP10 com os IPS que mais acederam ao site (no nosso caso so foram 3 IPs diferentes: local, outro tux e Proxy). Também apresenta um registo da duração das visitas. De notar que dependo do intervalo de tempo entre acessos, podemos apenas ter uma visita em múltiplos acessos. Apesar de termos centenas de acessos, apenas temos 8 visitas. Isto deve-se ao facto deste acessos terem sido com intervalos mais pequenos que o timeout definido no website.

No tipo de arquivos, podemos observar o tipo request que foram feitos. Os requestes predefinidos foram para a página php. No entanto, o acesso às ferramentas de análise é um pedido perl e por isso também é registado. As imagens png são extras das páginas das ferramentas de análise.

No fim da figura observamos os URL mais acedidos da nossa webpage, neste caso o domínio principal tal como seria de esperar.

3.3 Comparação

Em geral, as duas ferramentas são muito similares. Apresentam ambas essencialmente a mesma informação em gráficos bastante similares. Foi feita uma análise entre os dados obtidos pelas duas ferramentas para o mesmo site / registo de logs. Detetamos uma diferença com o número de acessos, que se prende com os acessos feitos aos Webalizer e ao AWStats durante a realização do relatório.



Visitas domínios/países (Top 10) - Lista completa

Domínios/Países		Páginas	Hits	Bytes
? Desconhecido	ip	389	428	831.28 KB
Outros visitantes		0	0	0

Hosts (Top 10) - Lista completa - Última visita - Endereço IP não resolvido

Hosts : 1 Conhecido(a)(s), 1 Desconhecido (ip não resolvido) 2 Visitantes únicos	Páginas	Hits	Bytes	Última visita
192.168.109.11	341	380	697.70 KB	14 Mar 2021 - 16:30
tux14	48	48	133.58 KB	14 Mar 2021 - 16:00

Motores/Spiders (Top 10) - Lista completa - Última visita

0 Motores visitantes*	Hits	Bytes	Última visita

* Robots mostrados aqui deram 'hits' ou tráfego "não visto" pelos visitantes, por isso não serão incluídos em outros gráficos.

Duração da visita

Numero de visitas: 8 - Med: 2,502 s	Numero de visitas	Percentagem
0s-30s	1	12.5 %
30s-2mn		
2mn-5mn		
5mn-15mn	1	12.5 %
15mn-30mn		
30mn-1h		
1h+	4	50 %
Desconhecido	2	25 %

Tipos de Arquivos

Tipos de Arquivos		Hits	Percentagem	Bytes	Percentagem
php	Dynamic PHP Script file	374	87.3 %	736.15 KB	88.5 %
png	Image	38	8.8 %	31.96 KB	3.8 %
pl	Dynamic Perl Script file	14	3.2 %	59.48 KB	7.1 %
gif	Image	1	0.2 %	430 Bytes	0 %
html	HTML or XML static page	1	0.2 %	3.27 KB	0.3 %

Downloads (Top 10) - Lista completa

Downloads: 0	Hits	206 Hits	Bytes	Dimensão média

Páginas/URL (Top 10) - Lista completa - Páginas de entrada - Sair

4 paginas diferentes	Acesso	Dimensão média	Páginas de entrada	Sair
/	369	1.98 KB	8	6
/awstats/awstats.pl	14	4.25 KB		
/webalizer/	5	1.30 KB		
/webalizer /usage_202103.html	1	3.27 KB		

Figure 3.10: Outras informações do AWStats (site2)

Destacamos uma funcionalidades implementadas no AWStats. A primeira é a capacidade de atualizar a informação com um botão no site. Este botão dá *trigger* a uma atualização no *host*, que recompila a informação com os logs mais recentes entretanto gerados. No Webalizer é preciso fazer isso manualmente ou criar um Cronjob que o faça por nós.

Outra diferença tem a ver com a facilidade de utilização. O AWStats foi implementado nos dois websites ao mesmo tempo. Tínhamos assim acesso ás estatísticas dos dois websites ao mesmo tempo.

O Webalizer, por outro lado, requere uma nova configuração para cada um dos sites. Só é portanto possível ver as estatísticas de um site de cada vez.

O AWStats também apresenta a duração das visitas, informação que o Webalizer não tem disponível.

Por fim, o AWStats apresenta uma interface mais atrativa, fazendo uso de gráficos mais modernos, que torna a página mais apelativa. Isto, no entanto, não é uma condicionante pois a funcionalidade e a documentação são os principais pontos a discutir.

A título de nota, o AWStats fornece uma análise de possíveis acessos feitos por um bot, informação que apresenta no início da página. Dado que 90% dos acessos foram feitos de forma autónoma e periódica e o AWStats não detetou quase nenhum, pode-se afirmar que o algoritmo precisa de algum trabalho.

Importante também referir que os logs do Squid também foram acedidos, onde podíamos observar o Ip de origem, neste caso o tux14, e os IP de destino, os 2 websites. O logs não foram analisados pois não foi possível instalar nenhuma das ferramentas de análise na máquina que continha o Squid. No entanto, no fase inicial de teste, conseguimos observar claramente o logging do Squid e do Apache os estes eram coincidentes tal como esperado.

4 Conclusão

Em suma, ambas as ferramentas são muito poderosas e apresentação uma boa análise estatística dos logs gerados. Pelos motivos acima referenciados, o AWStats aparenta ser a melhor solução, principalmente pela capacidade de servir vários websites ao mesmo tempo e de puder dar *reload* à informação diretamente da página. O AWStats também apresenta mais informação que Webalizer que o torna uma ferramenta mais poderosa.

O Webalizer por outro lado apresenta uma setup e uma interface mais simples, que o torna uma boa ferramenta para testes menos intensivos.

Existem outras soluções para análise de logs como é o caso de W3Perl. Uma desvantagem do W3Perl é o facto de ter deixado de receber suporte em 2015, o que torna a ferramenta obsoleta.

Um aspeto que gostávamos de ter melhorado neste trabalho é a geração de logs. Não fomos muito criativos na geração destes logs e eles não tem uma extensão temporal em que os gráficos apresentação sejam mais interessantes. Podíamos ter sido criados gráficos a simular uma utilização mais realista com utilização mais intensiva a meio da manhã por parte de um host e mais intensiva de tarde por parte de outro. Conseguimos no entanto aprender a utilizar estas ferramentas para analisar o impacto nos recursos de infraestrutura que alojar um site apresenta. Adquirimos também conhecimento sobre o funcionamento de um proxy e qual a sua utilidade em contexto real, especialmente neste trabalho onde a sua utilização foi especialmente necessária para podermos observar os resultados obtidos.