Edinburgh Napier UNIVERSITY

CSN09101
Networked Services
Week 10: Using Apache

Module Leader: Dr Gordon Russell
Lecturers: G. Russell

Edinburgh Napier UNIVERSITY

This lecture

- Apache Basic Authentication
- Log Analysis
- Security Issues
- Discussions

Edinburgh Napier
UNIVERSITY

Basic Authentication

Edinburgh Napier UNIVERSITY

Basic Authentication

- Often you might want simple usernames and passwords to control access you parts of a website.
- There are many approaches for this.
- The easiest way is to use Basic Authentication.
- This, when required, asks the browser to ask you for a username and password for accessing protected pages.
- The username and password is sent as clear text for every page request made by the browser.

Edinburgh Napier UNIVERSITY

.htaccess

- The best way to control basic authentication is via an .htaccess file in the directory to protect.
- To allow this the <directory> definition which includes the directory to be protected must have

  AllowOveride AuthConfig

Edinburgh Napier UNIVERSITY

Building a Password File

- You have to create a file with usernames and passwords.
- It is a good idea if this file is not one which someone can access via a URL.

> htpasswd –c /home/gordon/password andrew

New Password: *******

Retype New Password: *******

Adding password for user andrew.

-c is only the first time running the command, as this creates the file too. Miss out –c after the first run.

.htaccess

```
AuthType Basic
AuthName "Restricted Files"
AuthUserFile /home/gordon/password
Require user andrew
```

- Authtype Digest
  - This is another option, which requests the passwords in an encrypted format. It is not as widely supported as Basic.

Edinburgh Napier UNIVERSITY

The password file

- The password file created is just a text file.
- As a text file it does not scale well…
  - As more users are added the file gets bigger.
  - On every page request the file has to be parsed again.
- There are other formats available using hashed files (either db or dbm). These are faster to access but more complex to manage.

Edinburgh Napier UNIVERSITY

Any valid user

Require user andrew

- Can be changed to

Require valid-user

- In this way any user in the password file can access the directory.

Edinburgh Napier UNIVERSITY

Groups

- Just as in passwd users are also in groups, you can use the same idea for apache.
- Create a plain text file with the following format:

Groupname: user1 user2 user3 …

- If users gordon and andrew exists, and you want them to be known as group staff…

staff: gordon andrew

Edinburgh Napier UNIVERSITY

Add to .htaccess

AuthType Basic
AuthName "By Invitation Only"
AuthUserFile /home/gordon/password
AuthGroupFile /home/gordon/groups
Require group staff

Edinburgh Napier UNIVERSITY

Basic Auth Problems

- Its simple protection.
- Passwords in the clear.
- Every request need the password file lookup
- Large numbers of users difficult to manage
- Not a good idea for commercial systems
  - Yet some big sites use it!
- However, users recognise it and understand it.

Edinburgh Napier UNIVERSITY

Control by IP

- .htaccess can offer more control than just Basic Authentication.
- You can also restrict access to directories by IP.
- To do this you need to use
  - Order – read deny then allow or vice versa
  - Allow from – allow this match to access
  - Deny from – stop this match

Edinburgh Napier UNIVERSITY

Example

- Stop 10.0.0.1 accessing a directory…
- Edit the .htaccess in that directory:

order allow,deny
allow from all
deny from 10.0.0.1

Edinburgh Napier UNIVERSITY

Order is important

order allow,deny
allow from all
deny from 10.0.0.1

- This is identical to:

order allow,deny
deny from 10.0.0.1
allow from all

Edinburgh Napier UNIVERSITY

Domain Names

- You want to block anyone from jim.com and bob.com:

order allow, deny
deny bob.com jim.com
allow all

Edinburgh Napier UNIVERSITY

Development site

- You want only 10.0.1.0/24 and 10.0.0.2 to access the site:

order deny, allow
deny 10.0.1.0/24
deny 10.0.0.2
allow all

Edinburgh Napier
UNIVERSITY

Log Analysis

Edinburgh Napier UNIVERSITY

Logs

- Apache produces two types of log files
  - Error Logs
  - Access Logs
- Error logs are useful for debugging
- Access logs are excellent for monitoring how your site is being used.
  - Fun for people who have hobby sites
  - Life or death if your business relies on the web site.

Edinburgh Napier UNIVERSITY

Where are the logs

- Normally they go to /var/log/httpd/access_log and error_log
- In a virtual host we set them to what we liked:

```
<VirtualHost>
…
    ErrorLog logs/gr-error_log
    CustomLog logs/gr-access_log combined
</VirtualHost>
```

Edinburgh Napier UNIVERSITY

Logging in /var/log/http access file

- The normally used log format is called “combined”.
- It contains significant amounts of information about each page request.
- Specifically, the log format is:

%h %l %u %t %r %>s %b Referrer UserAgent

Edinburgh Napier UNIVERSITY

%h %l %u %t %r %>s %b Referrer UserAgent

- h – IP of the client
- l – useless ident info
- u – username in basic authentication
- t – time of request
- r – the request itself
- s – The response code (e.g. 200 is a successful request)
- b – size of the response page
- Referrer – who the client things told it to come here
- User Agent – identification info of the browser

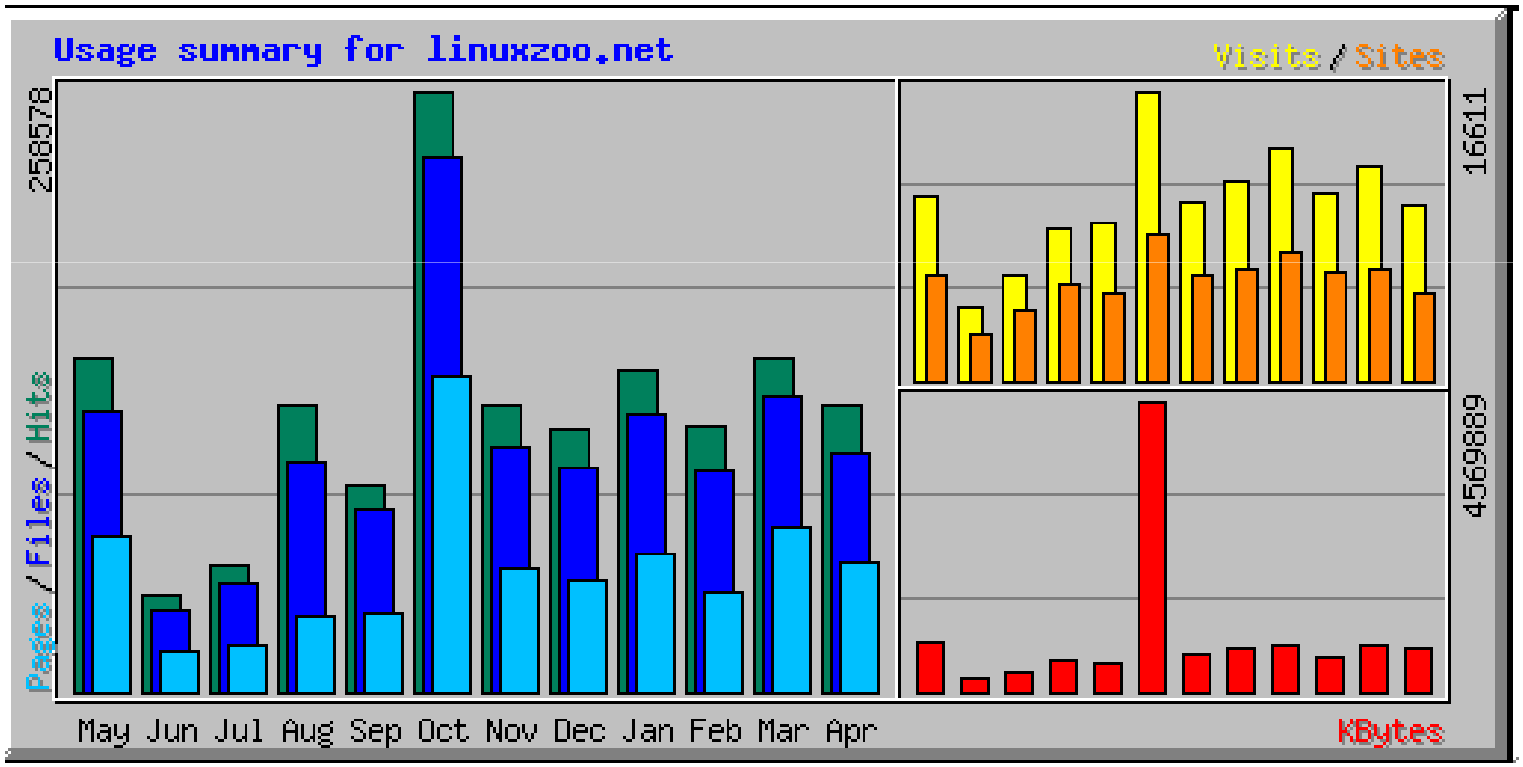
Edinburgh Napier UNIVERSITY

Analysing the log

- The log is useful in itself for checking the proper function of the server.
- However, traffic analysis is also valuable.
- There are a number of tools available to do this.
- One of the best free ones is webaliser.

Edinburgh Napier
UNIVERSITY

Webaliser Summary



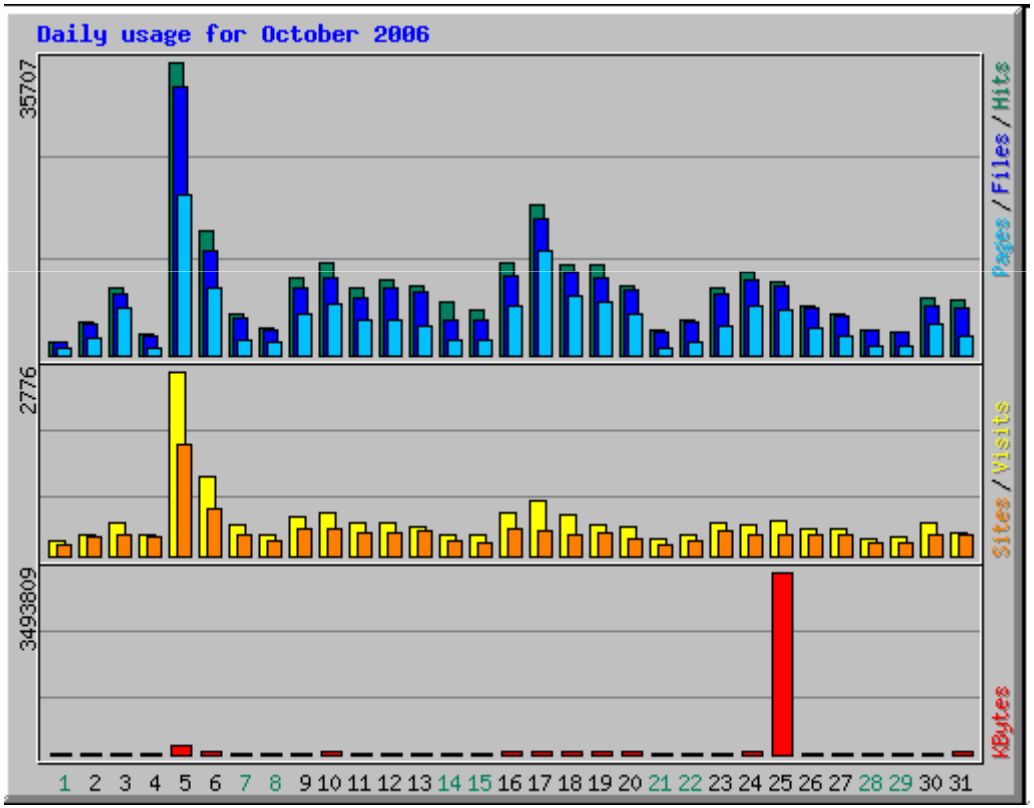
Edinburgh Napier UNIVERSITY

Analysis

- The summer is quiet for linuxzoo.
- Students are enthusiastic in October…
- After that it settles down to “kept busy”.

Edinburgh Napier UNIVERSITY

Per day activity – October

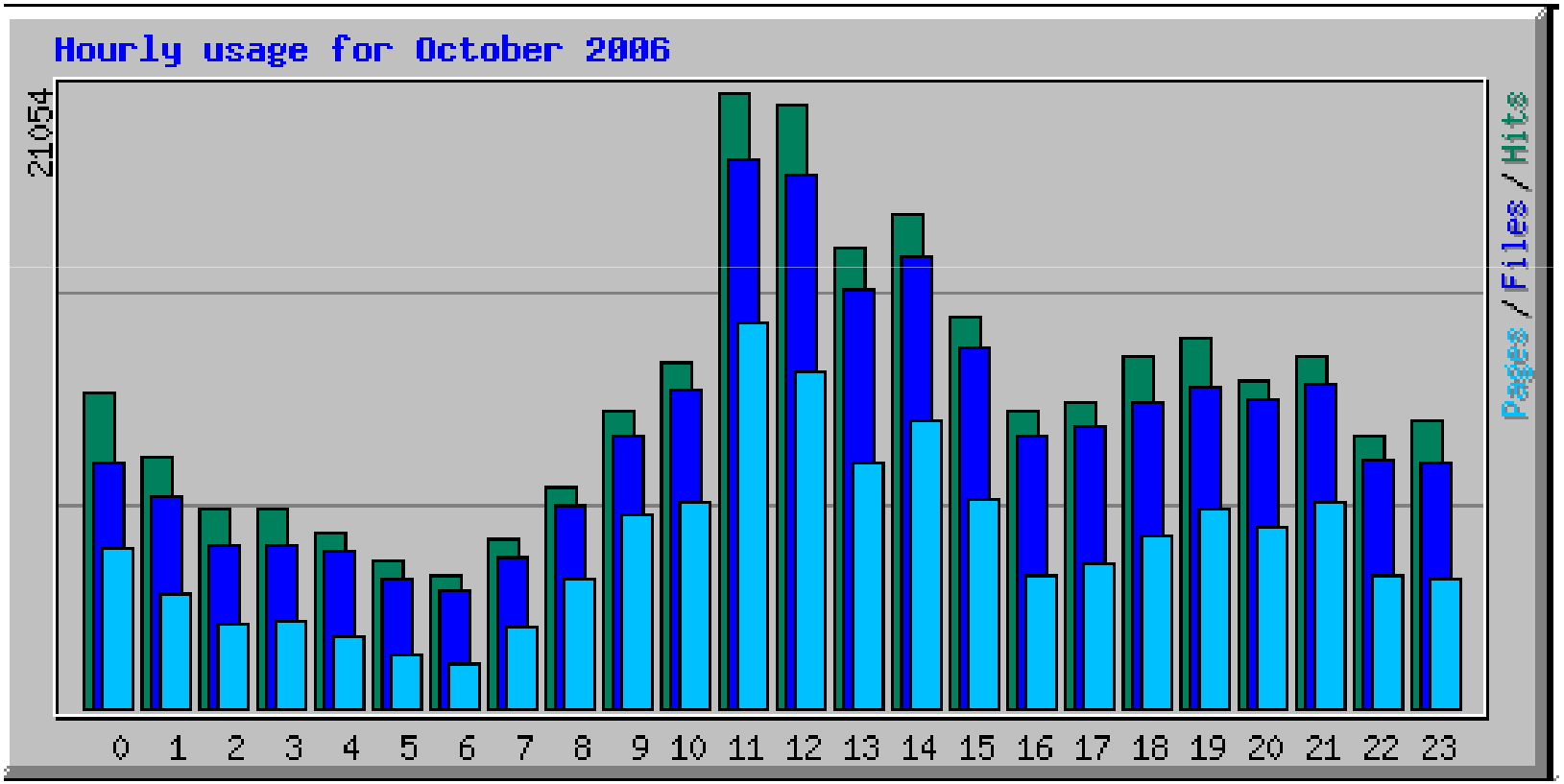


Edinburgh Napier UNIVERSITY

- I wonder which day was the first tutorial?
- Look at the 7 day oscillations. This is common in many web sites.
- Who stole all my web site data on the 25th?

Edinburgh Napier UNIVERSITY

Hour analysis – October



Edinburgh Napier UNIVERSITY

- Peak learning time (so they say) is 11am.
- Students here seem to like 9am-4pm.
- American students produce another bump later at night.

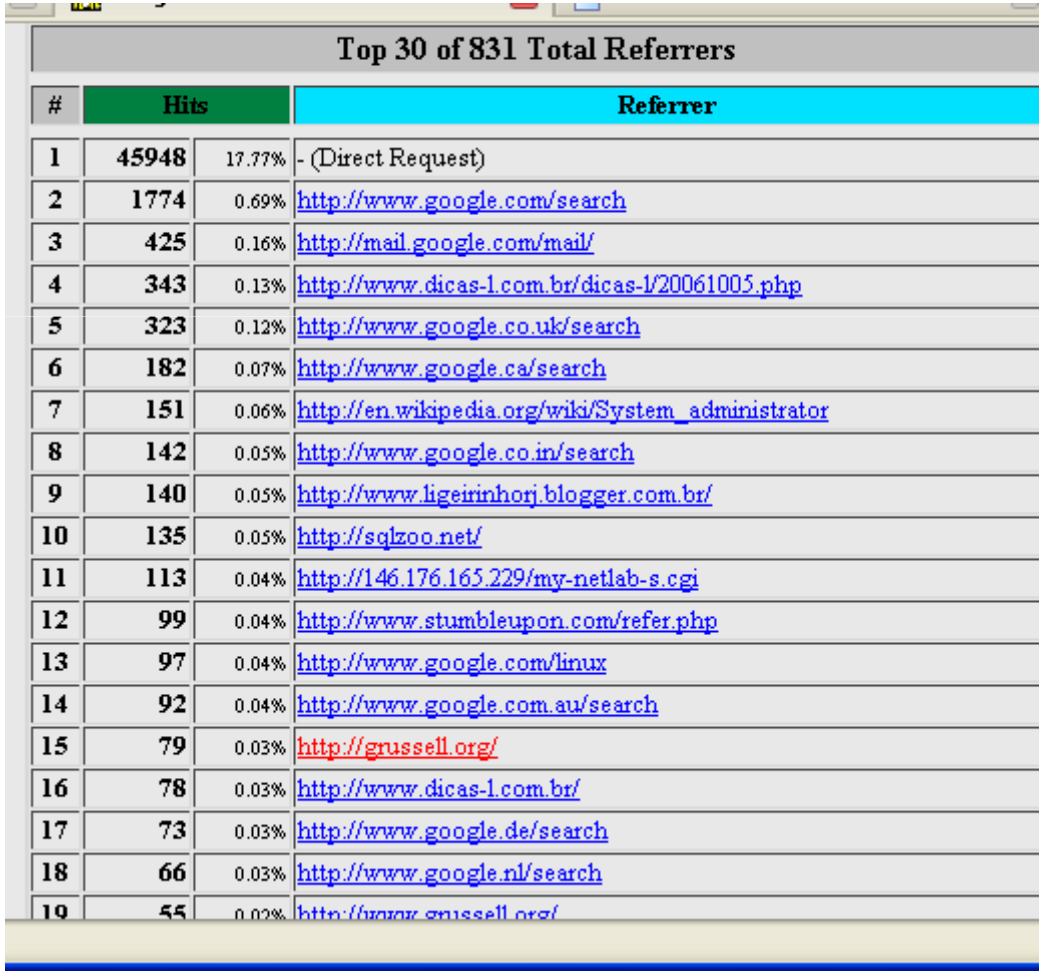
Edinburgh Napier UNIVERSITY

Users

#	Hits		Files		KBytes		Visits		Hostname
1	25319	9.79%	24605	10.69%	94474	2.07%	1051	6.33%	gtw-12.nhs.uk
2	16906	6.54%	16906	7.35%	14933	0.33%	1	0.01%	ds1081-020-069.nyc1.dsl.speakeasy.net
3	6400	2.48%	6381	2.77%	32857	0.72%	122	0.73%	200.182.252.4
4	6005	2.32%	5887	2.56%	24881	0.54%	245	1.47%	200.217.233.139
5	4903	1.90%	2230	0.97%	8144	0.18%	39	0.23%	mvx-200-196-55-148.mundivox.com
6	4506	1.74%	4506	1.96%	6580	0.14%	4	0.02%	200-221-128-3.corp.uolinc.com
7	3679	1.42%	9	0.00%	20	0.00%	2	0.01%	59.145.136.1
8	2728	1.06%	2726	1.18%	10462	0.23%	47	0.28%	59.163.124.54.static.vsnl.net.in
9	2690	1.04%	2590	1.13%	4733	0.10%	103	0.62%	58.68.28.66
10	2647	1.02%	2645	1.15%	11453	0.25%	193	1.16%	glenlivet.spc.eeng.liv.ac.uk
11	2479	0.96%	2455	1.07%	11568	0.25%	69	0.42%	mx2.queirozgalvao.com
12	2381	0.92%	2376	1.03%	50698	1.11%	1	0.01%	200.4.171.30
13	2156	0.83%	270	0.12%	2494	0.05%	6	0.04%	186.112-84-212.staticip.namesco.net
14	1657	0.64%	1657	0.72%	5821	0.13%	8	0.05%	80-41-249-174.dynamic.dsl.as9105.com
15	1606	0.62%	1601	0.70%	7577	0.17%	42	0.25%	146.176.242.35
16	1479	0.57%	347	0.15%	2545	0.06%	6	0.04%	r200-40-197-174.static.adinet.com.uy
17	1478	0.57%	1478	0.64%	0	0.00%	1476	8.89%	host.avidnetwork.com
18	1323	0.51%	1056	0.46%	12408	0.27%	3	0.02%	80-192-78-217.cable.ubr13.edin.blueyonder.co.uk
19	1157	0.45%	1155	0.50%	5840	0.13%	42	0.25%	146.176.242.54

Edinburgh Napier UNIVERSITY

Referrer Info



Top 30 of 831 Total Referrers

#	Hits		Referrer
1	45948	17.77%	- (Direct Request)
2	1774	0.69%	http://www.google.com/search
3	425	0.16%	http://mail.google.com/mail/
4	343	0.13%	http://www.dicas-l.com.br/dicas-l/20061005.php
5	323	0.12%	http://www.google.co.uk/search
6	182	0.07%	http://www.google.ca/search
7	151	0.06%	http://en.wikipedia.org/wiki/System_administrator
8	142	0.05%	http://www.google.co.in/search
9	140	0.05%	http://www.ligeirinhorj.blogger.com.br/
10	135	0.05%	http://sqlzoo.net/
11	113	0.04%	http://146.176.165.229/my-netlab-s.cgi
12	99	0.04%	http://www.stumbleupon.com/refer.php
13	97	0.04%	http://www.google.com/linux
14	92	0.04%	http://www.google.com.au/search
15	79	0.03%	http://grussell.org/
16	78	0.03%	http://www.dicas-l.com.br/
17	73	0.03%	http://www.google.de/search
18	66	0.03%	http://www.google.nl/search
19	55	0.02%	http://www.grussell.org/

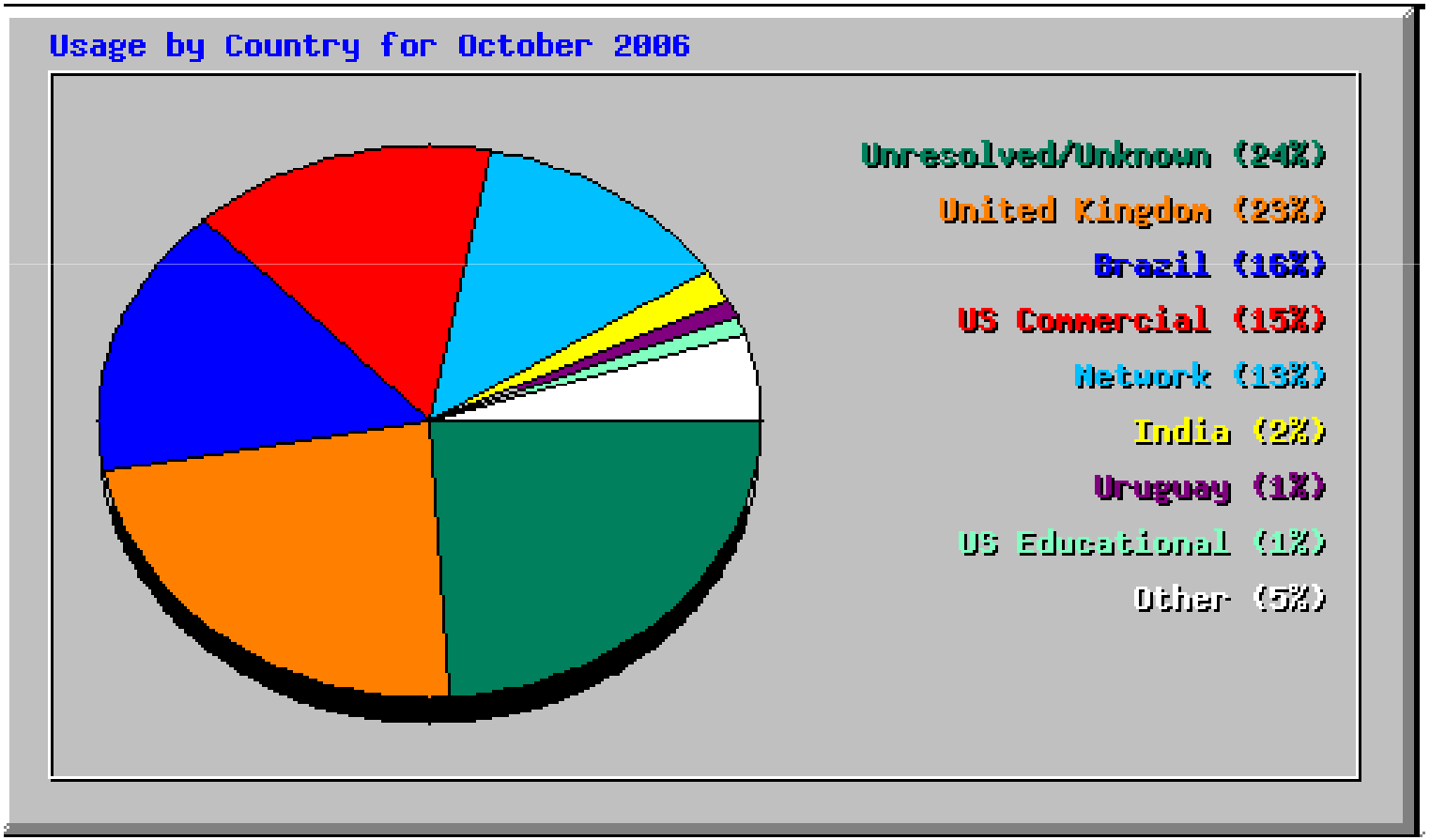
Edinburgh Napier UNIVERSITY

What search terms?

Top 20 of 1160 Total Search Strings			
#	Hits		Search String
1	176	8.17%	umask 022
2	68	3.16%	linuxzoo
3	46	2.13%	unix file structure
4	45	2.09%	apache2 authentication
5	31	1.44%	rndc.key
6	30	1.39%	apache basic authentication
7	27	1.25%	apache public_html
8	20	0.93%	generate rndc key
9	18	0.84%	linux zoo
10	18	0.84%	named service
11	16	0.74%	iptables -f
12	16	0.74%	mysql replicate
13	16	0.74%	symbolic link example
14	15	0.70%	apache2 basic authentication
15	15	0.70%	setting umask
16	14	0.65%	apache basic auth
17	13	0.60%	named forward
18	12	0.56%	rndc key

Edinburgh Napier UNIVERSITY

Where from?

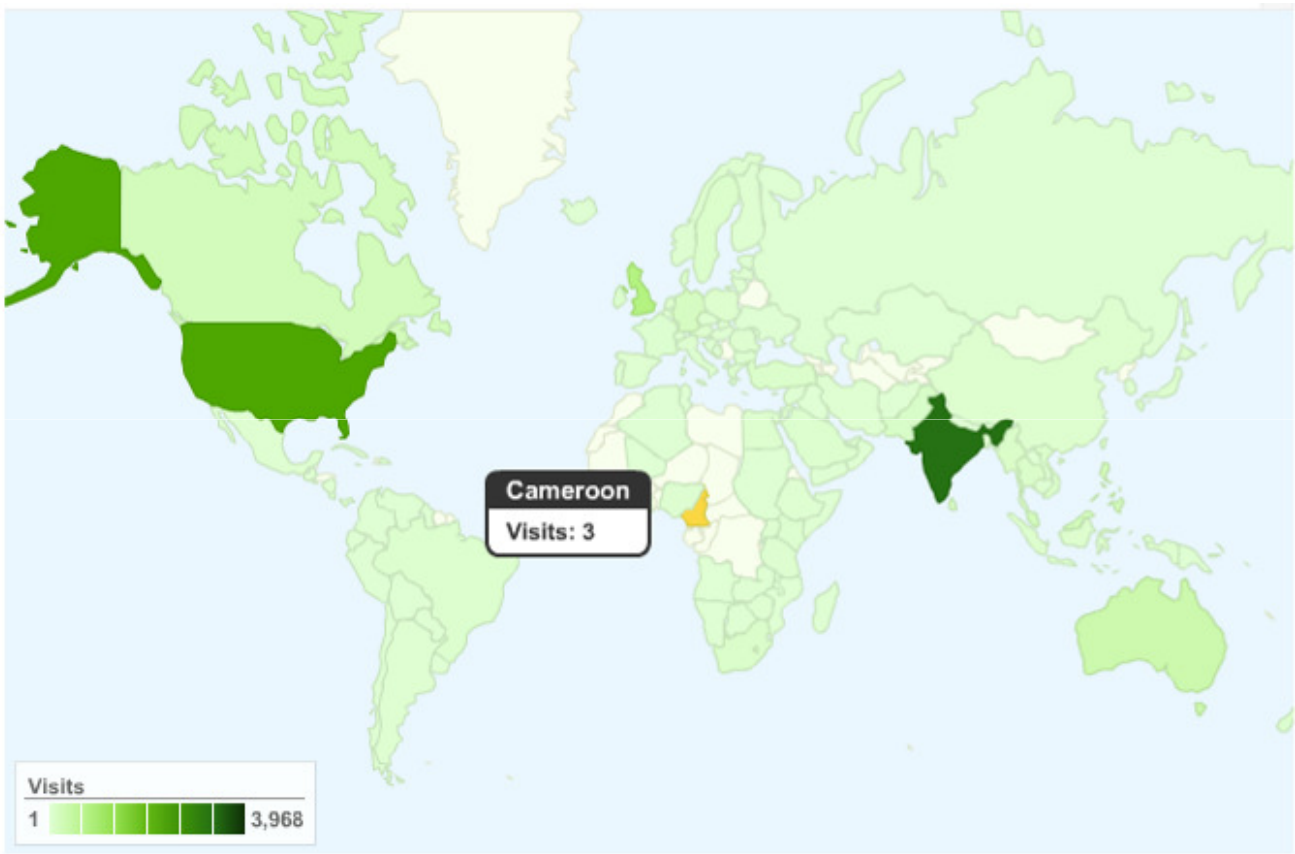


Edinburgh Napier UNIVERSITY

Google Analytics

- Another approach to web logging is to use JavaScript embedded in each web page.
- This does away with the need to access the web log.
  - Good if you don't have access!
- It does mean that
  - You only get logs where there is javascript switched on.
  - Each page is slowed by having extra stuff on it.
  - It's a little more complex.

Edinburgh Napier UNIVERSITY



11,998 visits came from 141 countries/territories

Edinburgh Napier UNIVERSITY

Logging Summary

- What is best?
- I have used both and have mixed feelings…
- Things to consider
  - Convenience
  - Reliability
  - Availability
  - Performance
  - Cost
  - Privacy
  - Complexity

Edinburgh Napier
UNIVERSITY

Apache Security

Edinburgh Napier UNIVERSITY

Security

- Hackers often consider a web server a good hacking target
- You should be very careful how apache is configured.
- The main problem is CGI scripts
  - CGI is a program which runs when you view a page.
  - Its output is sent back to the user's browser.
  - As it is an active process it can do permanent things to your server.

Edinburgh Napier UNIVERSITY

Simple CGI: who.cgi

```
#!/bin/sh
echo 'Content-Type: text/html; charset=ISO-8859-1'
echo
echo '<body><pre>'
whoami
env
echo '</pre></body>'
```

Edinburgh Napier UNIVERSITY

http://servername/who.cgi

apache SERVER_SIGNATURE=*Apache/2.0.51 (Fedora) Server at servername Port 80*
UNIQUE_ID=umn4CZKwogYAADNFYkcAAAAI
HTTP_KEEP_ALIVE=300
HTTP_USER_AGENT=Mozilla/5.0 (Windows; U; Windows NT 5.1; en-GB; rv:1.9.0.1) Gecko/2008070208 Firefox/3.0.1
SERVER_PORT=80
HTTP_HOST=servername
DOCUMENT_ROOT=/home/gordon/public_html

Edinburgh Napier UNIVERSITY

Issues

- This cgi program only prints.
- However, it could also delete things, or transfer data, copy passwords, etc.
- A hacker is rarely wanting distruction.
- Hackers want access! This requires either
  - Transferring hacking programs to the server
  - Copying files from the server (e.g. /etc/passwd).

Edinburgh Napier UNIVERSITY

Ideas

- Make sure apache runs as a user just for the server
  - The user “apache” is commonly used here.
  - In the httpd.conf, make sure there is:

    user apache

    group apache
- Hide the apache version number.
  - Might be useful if a hacker is searching for a buggy apache version.
  - In httpd.conf

    ServerSignature Off

    ServerTokens Prod

Edinburgh Napier UNIVERSITY

- Don’t allow apache to ever give pages from “/”

```
<Directory />
    Order Deny,Allow
    Deny from all
</Directory>
```

- Do you really need directory browsing?

  Options -Indexes

Edinburgh Napier UNIVERSITY

- The apache user should not own its conf files

  $ chown -R root:apache /etc/httpd

  $ chmod -R u=rwx,g=r,o-rwx /etc/httpd

- Do not allow apache to surf the web:

  $ iptables -A OUTPUT -m owner
  --uid-owner apache
  -m state --state NEW
  -j DROP

Edinburgh Napier
UNIVERSITY

Discussion

Edinburgh Napier UNIVERSITY

Discussion

- You want to secure apache so that all web requests can only use the characters a-z, “.”, and “/”. If they don’t then display the contents of “/noway.html”.

Edinburgh Napier UNIVERSITY

Discussion

- Here are some mock exam questions you should now be able to answer:

Edinburgh Napier UNIVERSITY

Question 1

- The following is an .htaccess file of a fictitious student on a student's web account.

```
Auth Type Basic
AuthTitle "Password Required"
AuthUserFile home/jim/.www-password
<Limit GET POST>
Required user jim
</Limit>
```

- Spot 4 errors

Edinburgh Napier UNIVERSITY

Question 2

- The following is an .htaccess file of a fictitious student on a student's web account.

```
AuthType Basic
AuthName "Password Required"
AuthUserFile /home/09006754/.www-password
<Limit GET POST>
Require user server_admin
</Limit>
```

- Provide the code to change the password for server_admin.

Edinburgh Napier UNIVERSITY

Question 3

- Below is a line from a webserver logfile:

157.55.18.25 - - [31/Aug/2011:12:48:04 +0100] "GET /robots.txt HTTP/1.1" 200 48 "-" "Mozilla/5.0 (compatible; bingbot/2.0; +http://www.bing.com/bingbot.htm)"

- What kind of request was this? Was this a successful request (i.e. was a document found)?