

Under attack

Gordon Woolf found his website had been fighting off raiders even as he slept

What does an attack on a web site look like?

The following is a single line extract from a fairly typical errors log on our web site. It occurred on January 10 this year in three hours with several lengthy gaps, so, for example, there were 10 attempts within one second, then a second or two gap and another burst, then several minutes before a further burst and a couple of lengthy gaps of almost an hour.

In total we had around a hundred attempts in this period and we gather from those who know a lot more than us that this indicated a very halfhearted attempt, probably by a robot of the kind which accounts for most of those window popups in individual firewalls such as ZoneAlarm.

In other words, it wasn't an attack on us, specifically, just by a program trolling for DNS addresses of computers that had left their door open.

These attempts all came from the one domain, and in between were errors from other domains with a message of "Broken pipe: client stopped connection before send mmap completed" which we take to mean that a legitimate web user had given up on trying to access our site.

A typical entry is:

```
[10/Jan/2002:16:06:15 -0004] [error] [client 130.94.19.147] File does not exist:
*/scripts/root.exe
```

where the asterisk indicates our internal directory structure which appears in the log file when someone enters our www address.

The attacker looked for root.exe in several folders, then sought files such as cmd.exe in a number of folders off the folder winnt, so it was obvious what system was being sought first -- Windows.

Then it looked for FrontPage folders such as _vti_bin/ which, despite our system not being a Windows host, we do have, but they are not activated, so our system's response to these was "unrecognized Frontpage request" in place of the usual "File does not exist".

In total this attacker made about 30 attempts to find cmd.exe.

Other attackers on the same day were more concerned with mail programs and were seeking our cgi-bin folder and, most frequently, the formmail.pl and formmail.cgi scripts. Our cgi-bin is not kept in the www folder, which makes it a little harder to gain access.

Presumably these were being sought as means of sending SPAM emails.

I'm not intending to promote fear among those with web sites, but it is a good reason for making sure that you can access logs from your web host and that you do occasionally glance through them to see what is happening while you aren't looking.

One reason for not promoting paranoia is that as well as the bad robots out there, there are many good ones -- those which patrol the net looking for web sites to promote the information they carry and update their entries on the search engines. There are others which gather general information -- like ever active census takers.

I've only scratched the surface of this aspect of running a website, and would

welcome more information from those who know more. Much of what is in the log files I do not yet understand yet, but these plain text files are also helpful in making our web site better, and picking up our mistakes.

For example, a whole series of error entries for the same tiny button graphic showed us that we'd put it in a different folder to where we thought it was. That was soon corrected.

And, a consistent entry for a web page directly off our root folder instead of in the subsidiary folder where it was, had us guessing that we'd got a mention on someone else's web page or in an article somewhere. That was soon solved by putting a file of the name being sought in the place where people were looking, and having it automatically transfer to the correct file. Much easier than trying to find the wrong entry and ask for a correction.

Log files can be hundreds of thousands of bytes, so they aren't to be read, just scanned, and there are automated programs to give you much of the useful information. However, a quick manual scan occasionally of the raw log files can show you things which can help make for a better website.

And the domain where the attacker seemed to live? It came up as belonging to Microsoft, so my guess would be that the attacker is good at hiding his tracks and spoofed an obvious candidate for blame.

-ends-