

How to determine Wordpress blogs with most spam on multiple blog hosting server

Author : admin



If you're a **hosting company** that is hosts **Joomla / Wordpress / [ModX](#) websites** (each) on *separate servers* and thus you end up with **servers hosting multiple Wordpress customer Blogs only**, lets say (**100+ WP blogs per host**) soon your **MySQL blogs databases** will be full (overfilled) with **spam comments**. **Blogs with multitude of spam comments** reduces the Wordpress site attractiveness, takes useless disk space, makes wp databases hard to backup and slowing drastically the SQL server.

As our duty as system administrators is to **keep the servers optimized (improve performance)** and **prevent spam-bots to hammer your Linux servers**, its is always a good idea to **keep an eye on which hosted blogs attract more spammers** and cause **server overheads** and **bad hardware optimization**.

Wordpress blogs keeps logged comments under **database_name.wp_comments (table)** thus *the quickest way to find out blogs with largest comments tables* is to use Linux's **find** command and **print** out only comments tables larger than set size.

Here is how:

```
find /var/lib/mysql/ -type f -size +1024k -name "'*_comments.MYD'" -exec ls -lh {} ; | awk '{
print $9 ": " $5 }'
```

`/var/lib/mysql/funny-blog/wp_comments.MYD: 15,7M`

`/var/lib/mysql/wordblogger/wp_comments.MYD: 5,3M`

`/var/lib/mysql/loveblog/wp_comments.MYD: 50,5M`

...

A comments database of 1MB means about at least 500+ comments, thus the blog **loveblog's** **wp_comments.MYD = 50,5** Mbs contains probably about **10000!** comments and should be definitely checked in a browser, if its overfilled with spam because of **bad anti-spam policy** or missing **currently best wordpress spam catcher plugin Akismet**. In cases of lack of client to protect his spam you can write quickly a script to auto mail him and ask him kindly to **check / fix his blog spam**.

In some cases it is useful to write a few liners bash script to automatically **disable users with extraordinary blog spam comments databases** (quickest way to do it is to *move users blog data under quarantine directory* and adding a Blog Suspended static html webpage with text like "**Please contact support for more info**").

1024k find arguments is 1MB, on a big hosted blogs this might be low and you might want to use (100Mb) = 102400kbytes.

You should note that *_comments.MYD in above find cmd is because though standardly wordpress sets wp_ as a prefix to its created skele table structures it is not always the case.

In above command example find looks for spam comments in **/var/lib/mysql** (because this is a Debian Linux server), however on other MySQL custom installs, it might be in another dir i.e. **/usr/local/mysql/data** etc.

It is useful to set the wp_comments statistics output to execute at least once a day as a cronjob:

```
crontab -u root -e 00 24 * * * /usr/sbin/check_spammed_blogs.sh
```

```
vim /usr/check_soammed_blogs.sh
```

Set a script like:

#!/bin/sh

*find /var/lib/mysql -type f -size +1024k -name "*_comments.MYD" -exec ls -lh {} \; | awk '{ print \$9 " " \$5 }' | tee -a /var/log/logs_with_most_spam_comments.log*

Though above commands is to run on GNU / Linux, for Windows servers based hosting you can install [GNUWin tools](#) and adapt above cmd using windows standard commands or **PowerShell** to do the same. Finally you can might want to use some other [SQL script to clear blogs with enormously large tables from spam or clear all unapproved spam comments](#)