Database installation notes

From SangerWiki

This page documents the process of installing the databases that are needed to run the Pfam website.

Contents

- 1 Assumptions
- 2 MySQL Config
    - 2.1 Setting up User Accounts
    - 2.2 Getting and installing the Pfam MySQL databases

Assumptions

We are going to assume that you already have MySQL v.5.0 (or greater) installed and that you have admin access to the database. If not, you may want to ask your database administrator for these privileges (or ask them nicely to do the database installation for you).

MySQL Config

We now store all information displayed on the website in the database, so there are no external file dependencies any more. As a result, all alignments are stored as BLOB (binary large object) which can be quite big. To get around this problem, you need to set the `max_allowed_packet` size to be greater than the default 1MB.

The best way to do this can vary, but usually when the mysql daemon is started you pass it a config file (commonly called as my.cnf). To alter the default packet size add the following line to the config file:

```
max_allowed_packet=16M
```

This will set the maximum packet size to be above what we need now and in the future. For this change to take effect, you will need to restart the mysql daemon.

Setting up User Accounts

For the new pfam database we use two databases. One database contains all of the Pfam data (which will be referred to as pfam_22_0), the other database is relatively small and acts as a broker for the offline jobs (referred to as web_user). We recommend 2 users for each database, one with only the essential privileges and the other for maintenance of the databases with all privileges.

Log in to the database as root (assuming that the database is local).

```
shell% mysql -root -p mysql
```

Generate the following user accounts for the Pfam database:

```
GRANT ALL on pfam_22_0.* to pfamwebadmin@127.0.0.1 IDENTIFIED BY "password";
GRANT SELECT on pfam_22_0.* to pfamweb@127.0.0.1 IDENTIFIED BY "FinallyAnotherPassword";
```

And the web_user database:

```
GRANT ALL on web_user.* to webuseradmin@127.0.0.1 IDENTIFIED BY "anotherPassword";
GRANT SELECT,INSERT,UPDATE on web_user.* to webuser@127.0.0.1 IDENTIFIED BY "yetAnotherPassword";
```

Test the accounts for connection:

```
shell% mysql -u pfamwebadmin -p pfam_22_0
Enter password: password
...
mysql> exit
Bye
shell% mysql -u pfamweb -p pfam_22_0
Enter password: FinallyAnotherPassword
...
mysql>
```

Getting and installing the Pfam MySQL databases

The current Pfam 22.0 database requires about 35GB when installed. You are going to need about three times this amount of space for the installation, as you will have to downloaded the files and the mysql log will probably capture the upload as well. There might be crafty ways around this, e.g. you could import the data table by table and clear the logs to reduce the overhead.

At the moment, there is a dump of the database in Rob Finn's ftp site. For users outside the Pfam consortium, this will need to be changed. Again, it is assumed that you have a way of getting the dump onto you local machine. For example, change to a scratch directory and do the following:

```
shell% wget ftp://ftp.sanger.ac.uk/pub/rdf/newWebsite.tar.gz
...
shell% tar zxvf newWebsite.tar.gz
shell% cd new_website
```

then to install the pfam_22_0 database, run something like the following bash commands:

```
bash% for t in `ls`
do
  echo "Uploading $t"
  cat $t | mysql -u pfamwebadmin -ppassword pfam_22_0
done
```

To install the web_user database do much the same.

```
shell% wget ftp://ftp.sanger.ac.uk/pub/rdf/newWebUser.sql
shell% cat newWebUser.sql | mysql -u webuseradmin -ppassword web_user
```

At each Pfam release, only the 'Pfam' database will need updating.

If you reached this point without any errors, then (in theory) you are good to go!

Retrieved from "http://scratchy.internal.sanger.ac.uk/wiki/index.php/Database_installation_notes"

Categories: PfamGroup | PfamWeb | PfamWebNotes

- This page was last modified 11:32, 7 September 2007.