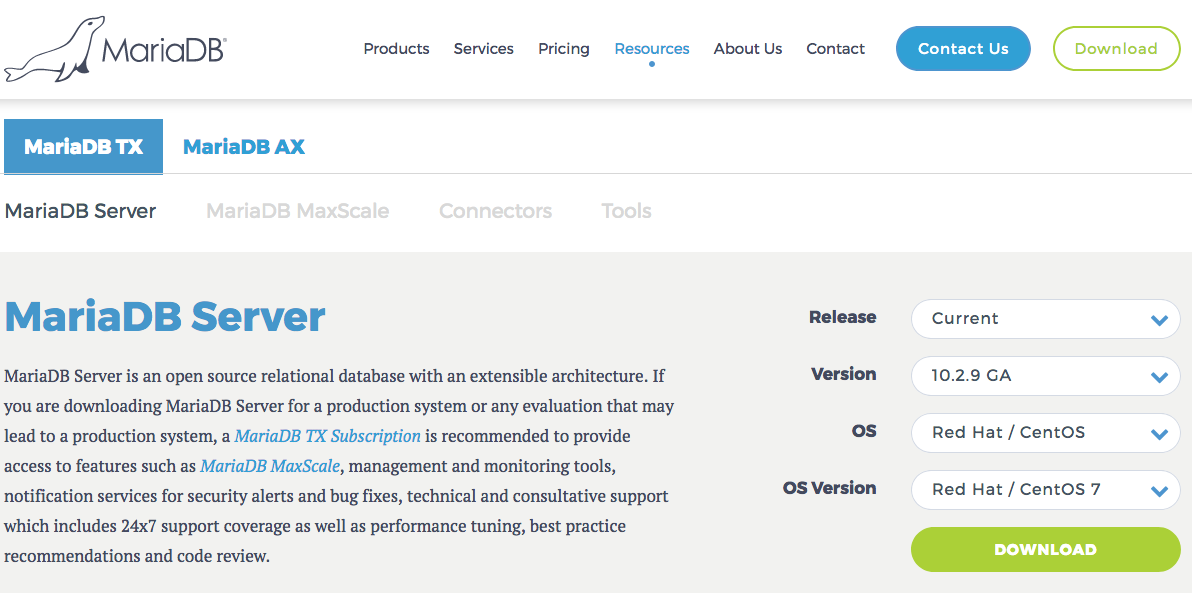
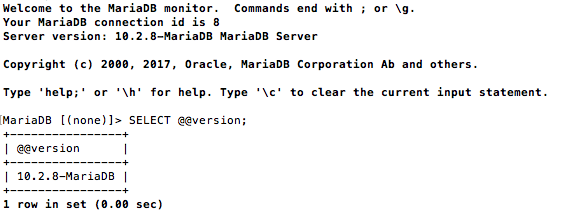
**Creating a MariaDB database with EAN Database content**

This short article describes how to create a database from the EAN downloadable content using MariaDB format as a target. Technically you will get ALL available data from EAN using this solution and you can use it as a central repository for a Database Server. This article focusses in a UNIX format server using the Linux Operating System. Please use the SQLite solution (in the documentation website) if you need just a simpler desktop solution.

First you will need to make sure you are using the latest MariaDB database engine and that it is up and running as per your particular Linux environment. You can download the latest version from its web site at <https://mariadb.com/>



We should have the **mysql** (now **mariadb**) command line utility added to your operating system path, we can then verify the engine version once we connected to the system running the database engine with:



In this case (a CENTOS 7 computer), we can see on the Server Version, that MariaDB version is 10.2.8 is running, so we are good to go! You can also query the DB engine with:

SE**LECT @@version;**

There are basically just 2 steps required to use MariaDB as your target database for imports. Let’s go over the required steps:

1. Create the empty structure of the database: You will need to run a .sql script we provide to create this empty structure.
2. Refresh the data: basically, download and refresh the data, you will probably need to run this step every time you need the latest information exposed by EAN. Due to the amount of data and depending on your Internet connection and speed of the machine you are trying to update; this process may take multiple hours.

Let’s examine each step, in details.

# Create empty database

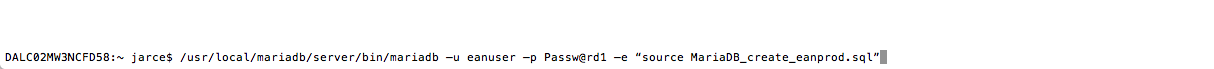
For any refresh project, you will need to create an empty database, that should create the storage area as well. The files are about 4 GB in size, once expanded it consumes around 7-8GB.

This step is easy to accomplish by using the mysql/mariadb command line utility, optionally you can use a graphical browser like Sequel Pro.

Using the script is quite easy, just run a command like:

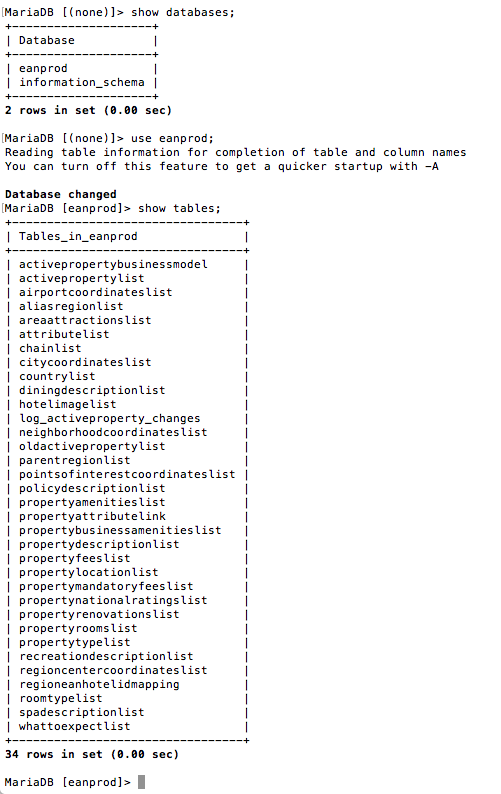
>mariadb –u eanuser –p Passw@rd1 –e “source MariaDB\_create\_eanprod.sql”

For our case, it will look something like:



In just a few seconds you should have the database created. If you need to cover other languages other than english you should extend the eanprod structure with the cooresponding tables. For example: we may need spanish and brazilian portuguese content, we will run the **extend\_eanprod**.sql script twice, replacing the XX\_xx with es\_es and then with pt\_br. Just edit the table names with the designated locale and run the extend version of the script.

We could do a technical verification of the database creation by using the same mariadb utility connecting to the newly created database and verifying its tables with the show tables command.



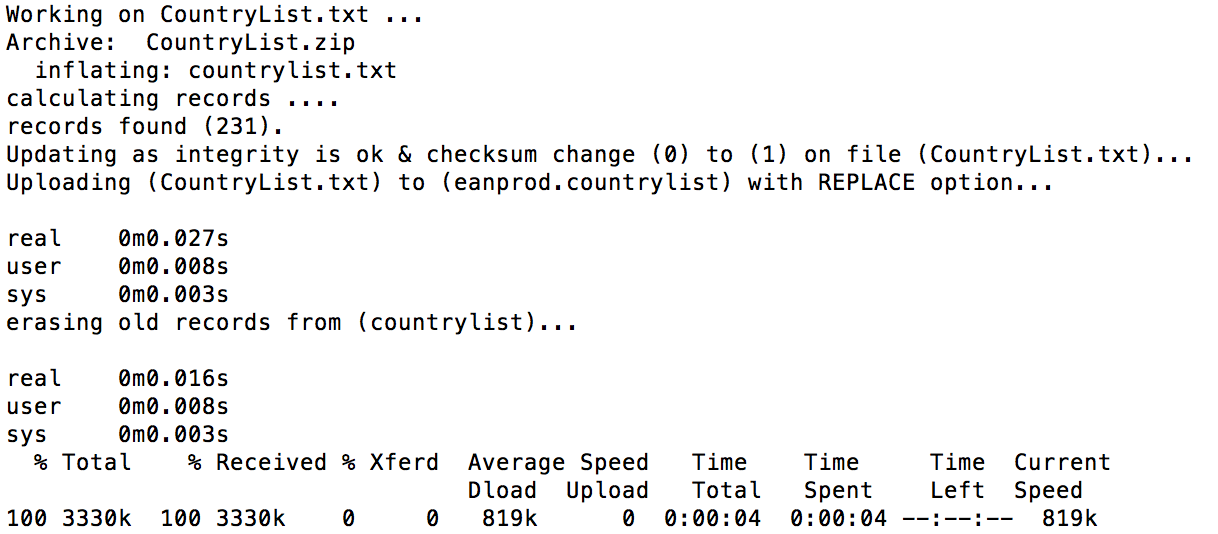
We now have an empty structure ready to receive the data from the EAN website. We first use the **show databases** command, then **use enaprod;** to use the **show tables** command. Remember to exit the mariadb utility with the exit; command, to finish.

# Refreshing the Data

Now that we have an empty structure or an existing one (even if it contains data) we can always pull the latest data by running the “refresh script”. Before we do so, we need to understand that the refresh script is created in the Linux platform under the bash shell.

***AFTER CHANGING AND ADJUSTING ALL DECLARATION OF VARIABLES IN THE .SH SCRIPT.*** The script is well documented and you can see all utilities that need to be presently installed for it to work (must Linux distributions have all commands already installed, but please double check). Change the script permissions to make it executable (we use the same eanuser user in this case) execute: EAN\_MariaDB\_refresh.sh

The first thing that it will do is to download the files from our site using the cURL command line utility. You should see something like:



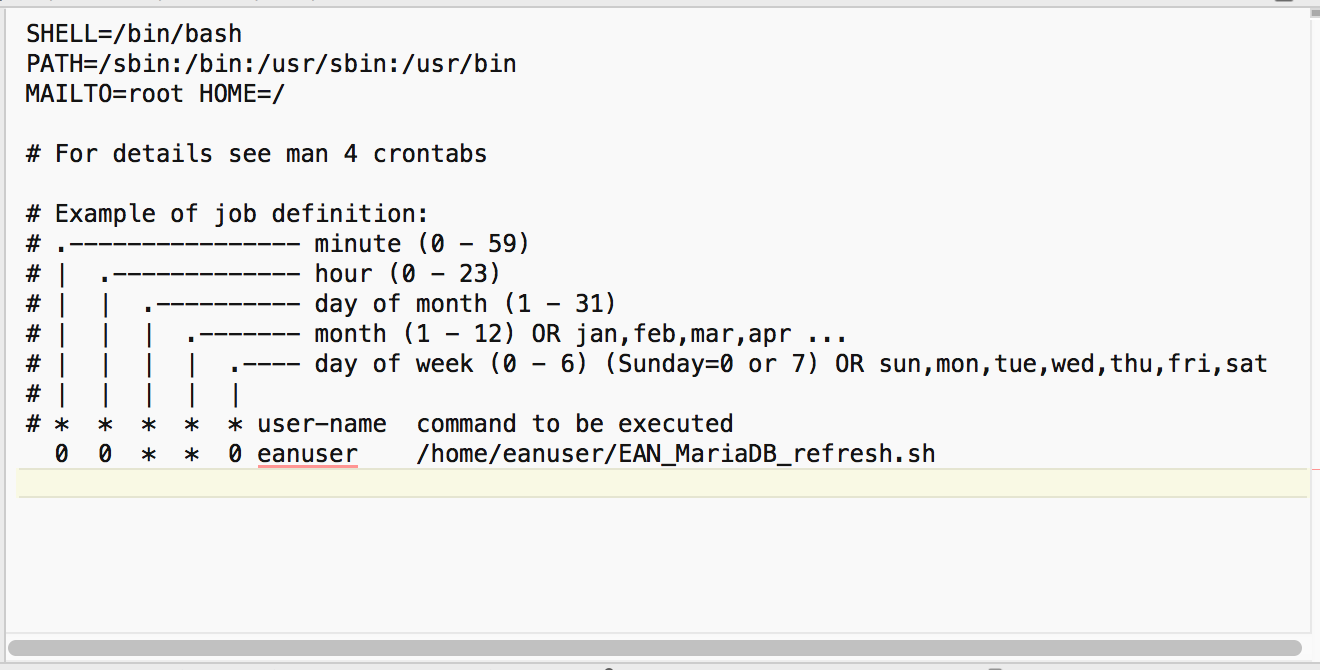
The script will tell you the download time and progress, you can also see if the file is different from the last time you downloaded. The “uploading...” line will be the indicator that the script is sending data to the database. Actual time will be affected by network connection and database engine performance but expect a few hours of work.

At the end of the script you will get a list of all tables with the amount of records, as well as the lines of records in each downloaded file.

# Schedule the refresh job

To schedule the refresh job you should use your operating system designed system. In case of Unix (so is the case with Linux) that is cron. To cron a job you just edit the crontab file. The documentation further states that the crontab command should be used to edit your crontab file, and you specifically edit the file with the **crontab -e** command.

Here is how our crontab file look:



In this case we configure the system to run at **00:00** (midnight) every Sunday, you can see we are using the **eanuser** to schedule the task and run it inside that directory. For our Linux version, we do not even need to stop/start the **crond** system as changes are pickup every minute by the engine. In our case the **crond** system email the **root** user when the job runs.

***We officially recommend running this task at least once a week.*** By-design the refresh script can run while the database and tables are in use, probably (by the massive amount of I/O operations) the performance will be impacted, so schedule it to run when your system is less in usage.

# Conclusion

In this paper, we discuss how-to create the database structure and refresh the data weekly to keep your system with the latest information of the EAN downloadable content. We focus on MariaDB as an example, but the same steps could be used with any other database engine, you will probably need to change how to upload the information and report on it.