

Forum:[Hack Forums](#)

Thread: [Database lookup](#)

Linus Torvalds - October 6, 2016, 2:07 pm

So I'm working on a database lookup service that will be completely free and will use MYSQL to search for all the information, however, I'm having issues inserting the databases unformatted information into MYSQL quickly. Does anyone have any ideas how to insert it quick? ?

LastSession V - October 6, 2016, 2:20 pm

No idea but good for you dude, very nice of you to do this for free :)

Abducted - October 6, 2016, 2:22 pm

Terminal and 'cut' are the last things you'll ever need in formatting strings, like db data in this case.

liumbiwe - October 6, 2016, 3:26 pm

So all your files are .txt , and do the data inside have some delimiters ?

MineSQL - October 6, 2016, 4:49 pm

Just insert the data with prepared statements if that's really the way you want to go. However, I would honestly use a tool that was made for stuff like this like elasticsearch

LSDoom - October 6, 2016, 4:54 pm

Don't use mysql. If you plan on adding loads of data (Billions of records) mysql will struggle. Not to mention the slow imports and hideous indexing times. Try using mongoDB instead.

I worked on a similar project for a while, MySQL was nothing but problems and we ended up quitting the project.

If you are still going for MySQL, try using something like navicat. It has an excellent import wizard which allows you to set up delimiters.

Specs - October 7, 2016, 5:07 am

You mean like LeakedSource? I attempted to make a site just like it using MySQL, however ended up stopping because it just wasn't cut out for the amount of data. Been reading about mongoDB so that might be the way to go.

Babur - October 7, 2016, 7:02 am

MariaDB can handle a couple of million records effectively. You might want to look at MongoDB, or use the cloud scalable service from AWS.

Linus Torvalds - October 7, 2016, 12:23 pm

Yeah many people have suggested it, I'm very uneducated on it however as I've only ever focused on mysql.

Scy7e - October 7, 2016, 12:29 pm

(Oct 7, 2016 12:23 PM)ripping Wrote: [\[CITE\]](#) Yeah many people have suggested it, I'm very uneducated on it however as I've only ever focused on mysql.

What a perfect opportunity and reason to learn something new.

Linus Torvalds - October 7, 2016, 12:39 pm

Indeed, do you have any good resources?

Scy7e - October 7, 2016, 1:37 pm

(Oct 7, 2016 12:39 PM)ripping Wrote: [\[CITE\]](#) Indeed, do you have any good resources?

Honestly, I'm not really a fan of MongoDB so I can't really give you any good advice regarding it, and I doubt you can get that much better results than with MySQL.

I don't want to discourage you, but this is definitely not an easy job. A database lookup is essentially useless unless you have hundreds of millions or billions of records and searching through that isn't simple. I'm not saying it's not doable, but you're gonna have to make compromises if you want to offer this for free, since it's gonna cost you.

If you want the users to be able to lookup by email, it's rather easy to optimize the DB, but if you want to be able to lookup by any keyword, things get complicated. You'll want to scale out, not up. Rather many cheaper computers than one high-end.

As for what DB should you use? It depends entirely on how you're going to do this and what's good enough for you. Maybe the way haveibeenpwned.com did it is good enough? They use Azure so don't expect it to be free: <https://www.troyhunt.com/working-with-154-million-records-on/>

Linus Torvalds - October 7, 2016, 2:05 pm

Yeah I'll have to find some way, I hope to not give up never the less

SIM-1407431 - October 8, 2016, 10:35 am

I once made a db lookup using GREP, however it wasn't the best solution, but it did the job.

Linus Torvalds - October 8, 2016, 11:09 am

Yeah if you want to be able to index billions of records and have more than one person grepping at one time, the server would probably die