Sphinx

1.   What is it?

Sphinx is an open source full text search server, designed from the ground up with performance, relevance and integration simplicity in mind.

It's written in C++ and works on Linux (RedHat, Ubuntu, etc), Windows, MacOS, Solaris, FreeBSD and a few other operating systems.

Sphinx lets you index and search data stored in an SQL database, NoSQL storage, or just. Queries can be batched or run on the fly.
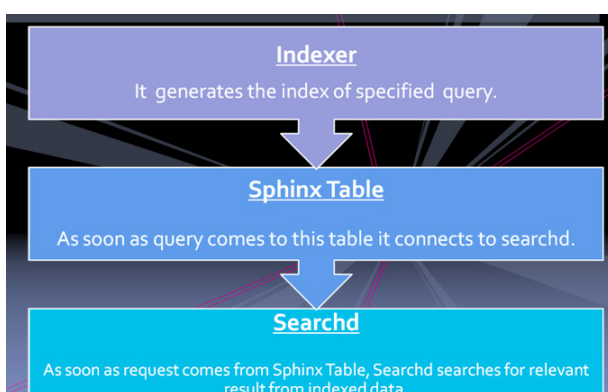

Main features are:

- Batch and Real-Time full-text indexes
- SQL database indexing
- Easy application integration
- Advanced full-text searching syntax
- Distributed searching

2.   How does it work?

Sphinx is available as stand-alone server for MariaDB and MySQL (using the database protocol or a proprietary API) or as a Storage Engine, known as SphinxSE.

Despite the name, SphinxSE does not actually store any data itself. It is actually a built-in client which allows the database server to talk to searchd, run search queries, and obtain

search results. All indexing and searching happens outside MariaDB or MySQL. This means that full-text search queries can be run as a normal SQL query, the result set can be joined or merged with result sets coming from other (MyISAM & InnoDB) tables.

SphinxSE is included in the MariaDB distribution from v5.2.22 onwards.

3.   When to use it?

- When you have a large number of documents
- When search speed is a strong requirement
- When you have to combine search results with other relational data
- When you want to use the standard mysql protocol instead of a proprietary API

4.   Resources

- Sphinx Storage Engine Manual: sphinxsearch.com/docs/current.html
- Enabling Sphinx in MariaDB: kb.askmonty.org/en/about-sphinxse/2941/
- Sphinx Conference Talks: sphinxsearch.com/info/talks/