Unified Index Comparison

The Search and List from Unified Index has support for multiple engines. While all of them offer the same general functionality and connect to various functionality such as the content search, PluginList, PluginCustomSearch and various others, they will have different performance characteristics and some may offer additional features.

As a general rule, the engine can simply be switched and the index rebuilt without any additional change to the configuration.

Overview

The unified index engines are:

- MariaDB / MySQL / Percona Server for MySQL Full Text Search **(is the default for 12.x onwards)**
  - *introduced in Tiki12*
  - Additional memory required
  - Fast indexing (can be 10 times faster than now removed Zend_Search_Lucene), slower/unstable query speed
  - No configuration required
  - Not customizable
  - Stored in tables in the database with a prefix of index_ and are stored in MyISAM even if the actual data is in InnoDB
- Elasticsearch
  - *introduced in Tiki12*
  - Independent Java server(s), horizontally scalable
  - Feature-rich
  - Fast indexing, fast/stable query speed, decent/good results
  - Typically, Elasticsearch is set up as a cluster on different servers than Tiki (or using a third-party service), but it is also possible to install on the same server.
  - Customizable
- Manticore Search, new in Tiki25
  - Feature-rich
  - Very fast
  - Written in C++ with Manticore Buddy in PHP
  - Customizable
  - Can be set up as a cluster
  - Requires small amounts of RAM. (compared to Elasticsearch)
    - It runs fine on a virtual machine with 1 GB of RAM
  - Great support for PHP
  - Can be installed without root access so shared hosting should be OK (They will need SSH though).
  - This is the default setup for WikiSuite once Tiki26 is released, and it is an option of the installer.

The system is designed for maintaining an autonomy vis à vis the engines. So more can be added later. No long-term data is stored in the indexes and it's fairly easy to switch from one to another. The next logical addition is OpenSearch. Please contact Marc Laporte if you have specific needs.

Limitations

MySQL/MariaDB

- Words with fewer than 3 or 4 characters will not be indexed unless the server configuration is modified. Variables name: ft_min_word_len and innodb_ft_min_token_size

- Comes with an extensive list of English stop words, preventing many queries from working.
- Can use a single index at a time. Depending on the query, performance can vary significantly.
- Several limitations on the number of columns and indexes it can contain. Complex sites with many different query patterns may hit those limitations.
- No support for field boosting, such as providing more relevance for hits on the title.
- There is a limitation on the number of tracker fields. The limitation is quite high (2000+), but when you hit it, you need to move to another engine because MySQL/MariaDB has a hard limit. It is not possible to know in advance the precise number of maximum fields because some tracker field types require more than one column.
  - It is possible to exclude some fields from the index to stay within the limit.
- How to search currency amounts likely produces bad results (to be tested)

Elasticsearch

- No longer Open Source
- Requires a dedicated environment to be installed and works better with multiple instances running in a cluster.
  - Requires Java and a lot of RAM/CPU
    ❝

    > *A machine with 64 GB of RAM is the ideal sweet spot, but 32 GB and 16 GB machines are also common.*
    > *Source:*
    > *https://www.elastic.co/guide/en/elasticsearch/guide/current/hardware.html#_memory*

Manticore

There is a hard limit of 256 full text fields per index. Additional fields will be slower.

Extra features

- Stored Search
  - Only supported by Elasticsearch and Manticore
- Faceted search (dynamic filters applicable on search results)
  - Only supported by Elasticsearch and Manticore
- Module More Like This
  - Only supported by Elasticsearch (Manticore on roadmap)
- Federated Search
  - Only supported by Elasticsearch (Manticore on roadmap)

Selection guidelines

Tiki 25 and before

- Small sites, simple functionality: MariaDB/MySQL Full Text Search
- Medium or large sites, advanced functionality: Elasticsearch

Tiki 26 and up

- Small sites, simple functionality: MySQL Full Text Search
- If you are already using Elasticsearch and are happy with it: Elasticsearch
- Medium or large sites, advanced functionality: Manticore

Speed comparison

- https://db-benchmarks.com/?cache=fast_avg&engines=elasticsearch%2Cmanticoresearch_rowwise%2Cmysql&tests=hn_small&memory=110000&queries=0%2C1%2C2%2C3%2C4%2C5%2C6%2C7%2C8%2C9%2C10%2C11%2C12%2C13%2C14%2C15%2C16%2C17%2C18%2C19%2C20%2C21%2C22%2C23%2C24%2C25%2C27

Differences of results between engines

index:compare-engines

Legacy

Zend_Search_Lucene (PHP Implementation) was introduced in Tiki7 and later removed Last version: Tiki21.

alias

- Search Engine Comparison