

Unified Index Comparison

The [Search and List from Unified Index](#) has support for multiple engines. While all of them offer the same 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:

- **MySQL Full Text Search (is the default for 12.x onwards)**
 - *introduced in [Tiki12](#)*
 - Within MySQL, additional memory required
 - Fast indexing (can be 10 times faster than `Zend_Search_Lucene`), slower/unstable query speed
 - No configuration required
 - Uncustomizable
- [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 ([This is the default setup for WikiSuite](#))
 - Customizable
- [Manticore](#) is coming soon
- `Zend_Search_Lucene` (PHP Implementation) **Removed: Last version [Tiki21](#)**
 - "introduced in [Tiki7](#), and will be removed at one point since the Zend Framework project leader wrote: "[This component has not been maintained since we split it from the ZF1 sources; we have no plans to make any further releases.](#)" and "[This component is no longer maintained; the last PHP version it was tested against is 5.3.](#)"
 - Complete PHP implementation
 - CPU and Disk I/O bound
 - Slowest indexing, stable query speed, decent results
 - Customizable
 - Requires file permission configuration (Apache needs to be able to write to the file)

Others may be added later

Limitations

The `Zend_Search_Lucene` implementations's primary limitation is indexing performance.

- It is heavily bound on I/O disk operations and indexing will consume important amounts of PHP memory.

- Query caching may cause the disk to use large amounts of space.
- When the search query is too wide and provides over 1000 documents (configurable), it will not fully explore all possibilities to keep ranking time reasonable.

The MySQL implementation has several limitations:

- Words with 3 or less characters will not be indexed unless the MySQL server configuration is modified.
- MySQL comes with an extensive list of English stop words, preventing many queries from working.
- MySQL can use a single index at a time. Depending on the query, performance can vary significantly.
- MySQL has 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 Zend_Search_Lucene 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 MySQL column.

Elasticsearch requires a dedicated environment to be installed and works better with multiple instances running. It does not have known limitations.

Extra features

- Facet search (dynamic filters applicable on search results)
 - Only supported by Elasticsearch
- More Like This
 - Only supported by Elasticsearch
- [Federated Search](#)
 - Only supported by Elasticsearch

Selection guidelines

- Small sites, simple functionality: MySQL Full Text Search
- Small/medium sites, advanced functionality: Zend_Search_Lucene (PHP Implementation)
- Dedicated environment: Elasticsearch

Engine-specific notes

Zend_Search_Lucene (PHP Implementation)

The default implementation is based on Zend_Search_Lucene, a PHP implementation of the Java Lucene index engine. The engine has no external dependencies and can run on all hosts. However, some configuration may be required to reach acceptable performance.

The time required to build a complete index will vary depending on the content of the site. As a reference,

- doc.tiki.org (this site), with over 1400 pages, reindexes in around 3 minutes
- themes.tiki.org, with some pages and several hundreds of forum posts, reindexes in around 15 seconds
- corporate intranets with several gigabytes of data in file galleries can take over an hour

Alexander Veremy, the author of the component provided [some insight on how to adjust the parameters](#).

MySQL Full Text Search

No notes.

Elasticsearch

No notes.

Differences of results between engines

A tool has been created to compare the results, and make sure all is OK:

https://gitlab.com/tikiwiki/tiki/-/merge_requests/940/

alias

-
- [Search Engine Comparison](#)