Abstract

In IT industry, one of the most important goals of the organization is to reduce overall cost of development. One way to achieve this is making the developers to write modular and reusable code organized in subroutines or packaged libraries. The difficult challenge is to search through code spanning across millions of lines used in various legacy projects. This paper explains about a framework that aims to develop a search engine which enables easy searching through large repositories of source code. It also goes further to achieve scalability and high performance through the deployment of search engine in a distributed system.

References

1. Mining modern repositories with elasticsearch
   Kononenko, O Baysal, R Holmes
5. https://www.elastic.co/products/elasticsearch

**Index Terms**

<table>
<thead>
<tr>
<th>Computer Science</th>
<th>Information Sciences</th>
</tr>
</thead>
</table>

**Keywords**

Solr, Distributed Search, Zookeeper