Apache solar

Introduction

Solr was created in 2004 as an in house project to add search capability for the CNET networks. It was donated to APACHE Software Foundation after CNET Networks decided to make it open source. After a number of versions of the Solr, with every version having enhancements, the latest version today is the Solr 5.0 which is a standalone application. Solr is an open source search platform written in JAVA. It has full text search capabilities that enable matching phrases, hit highlighting, real time indexing so that the user can see the content whenever he wants to see, no SQL features, extensible plugin architecture, dynamic clustering, database integration. Providing distributed search and index replication, Solr is highly scalable and fault tolerant.

Why Apache Solr is so Popular?

Although there are databases and frameworks such as HADOOP, Apache Solr has been on top of these due to its standalone noSQL store. Solr effectively serves as a data access layer for doing key value lookups as well as making the data fully indexed and searchable. Search is always the first requirement in many deployments, which is pushing many organizations to use the search engine like a noSQL store for that robustness and flexibility.

Another aspect that makes Apache SOLR popular is the governance model enforced by it to maintain accuracy of records. However Solr uses Apache Zookeeper to handle distribution and puts a governance model behind the data retrieval process to guarantee delivery. The main difference with NoSql data sources is they are optimized around lots of writes and they are constantly updating so that they can keep track of all the different searches and transactions done.

Solr has also addressed the deep paging problem because the deeper you go into results the slower the query will be by introducing the cursor. This enable the users to quickly navigate through pages.

Do users use Apache Solr?

Google is arguably the worlds largest open source company, that release not only countless lines of open source code but also hosts many open source projects. Despite these open source giants, why would users go ahead to use Solr? Solr is a cheap and excellent option to build a quick search experience. Some Companies are using solr due to its fast searching and indexing speeds, Flexible data model, great query API, Buit on top of proven technology (Lucene). It is easier to roll a service using Solr that can help improve the search quality for users and google and other giants like Microsoft and Oracle are adopting their ways of working as per the new use cases and data models.

Appropriateness and usefulness of documentation

The Apache Solr has a Wiki page where all the official documentation about every release can be found. Of particular it has the Solr reference guide which is published by the project after each minor release. The Wiki page cover all the topics for including installation and configuration, search and indexing, Solr Cloud , Advanced tools, Tips, Tricks and Use cases, Solr Clients, Operations and Production and user contributed content. The code is well maintained and readable. Open source softwares live on for many years, and it is very important that can be adopted by other professionals who can read and understand the code. Apache Solr having so many version since 2004, explains the readability and maintainability of the code.