# APACHE SOLR

A Quick Overview
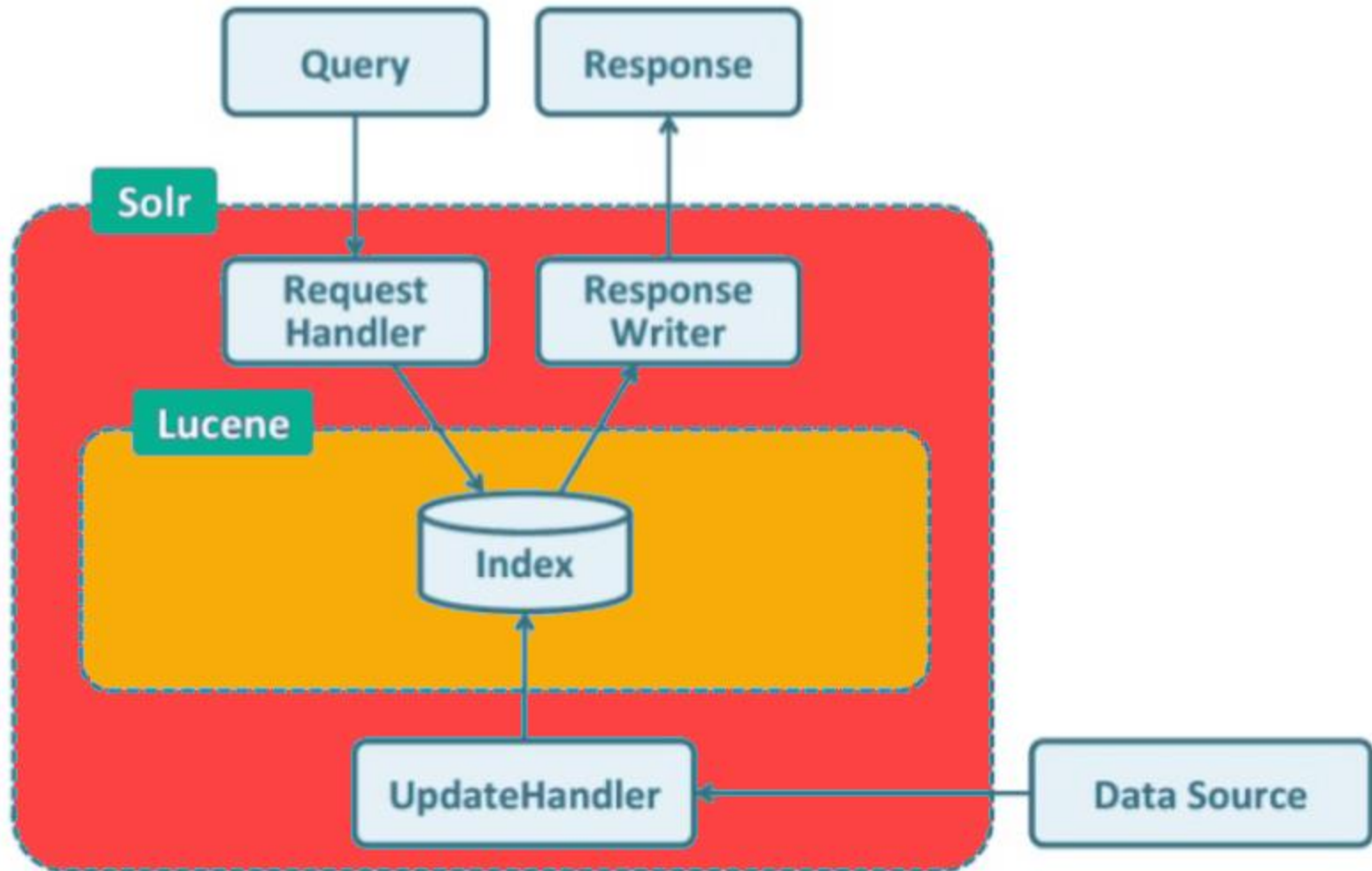
by  Suchitra Jayaprakash

# WHAT IS SOLR ?

- Open-source REST-API based Enterprise Real-time Search and Analytics Engine Server.

- Highly scalable search applications

- Ready to deploy

- Built using Apache Lucene Framework , Java

- Optimized to search large volumes of text-centric data

- Document-based NoSQL Data Store

- Main Functionality – Indexing & Searching

# SOLR FEATURES

- Full text search

- Faceted Navigation

- Spell check

- Hit highlighting

- Relevant result

- Recommendation

- Geo-Spatial Search

- supports Distributed and Cloud Technology

- Built in Authentication & authorisation

- Learning to rank

# SOLR ARCHITECTURE

# INSTALL SOLR  LOCALLY

- Java Runtime Environment version 8 or greater

- Download Solr from lucene site and unzip for installation.

- Solr Instance is an instance a Solr running in the JVM

- Start/Stop Apache solr
    - bin/solr start
    - bin/solr stop
-  Access the Solr Admin User Interface
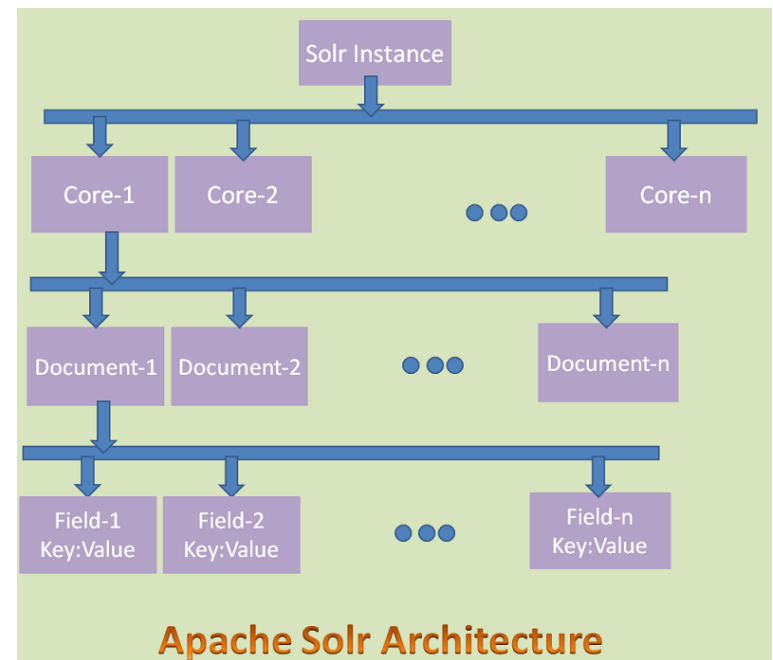    - http://localhost:8983/solr/
-

# CREATE CORE

- A Solr Core is a running instance of a Lucene index that contains all the Solr configuration files required to use it.

   Core = an instance of Lucene Index + Solr configuration

- We need to create a Solr Core to perform operations like indexing and analyzing

- solr create_core -c XXXX -p 8983

# INDEXING DATA

- Solr can ingest many types of files  like .csv .json, .xml, .html,.pdf
- We can add data to Solr index in
    - Solr Web Interface.
    - Client APIs like Java, Python, etc.
    - post tool.

- Apache Lucene Inverted Index technique to Index it's documents

- Use Admin console or POST
- command to upload document



**Apache Solr Architecture**

- java -Dc=XXXX  -Durl=http://localhost:8983/solr/XXXX/update/  -Dtype=application/xml -jar ../example/exampledocs/post.jar   "../example/films/films.xml"

# CONFIGURATION FILES

- solr.xml - server instance configurations

- core.properties - core configurations such as names, locations and files in the core

- conf/solrconfig.xml - core configurations for field guessing, directories, query settings, spell checking, keyword highlighting and query response formats

- conf/managed-schema - core configurations for field processing

# SEARCH IN ADMIN UI

- Open Solr Admin Console: http://localhost:8983/solr/

- Select the core module

- Click on "Query" from the left navigation

- Enter search query in  q text box & df (default field)

- Click on "Execute Query" button

- It retrieves matching document from  selected core

# SEARCH ….

- You can also do search using rest api call

  - http://localhost:8983/solr/<corename>/select?indent=on&q=*:*&wt=json

  - http://localhost:8983/solr/<corename>/select?q=black&mlt=true&mlt.fl=genre&mlt.mindf=1&mlt.mintf=1&fl=id,score&df=genre

  - http://localhost:8983/solr/<corename>/spell?spellcheck=true&qt=spellchecker&spellcheck.accuracy=0.5&spellcheck.collate=true&q=Cime

# THANK YOU