

Week 8 Assignment: Diving into Apache Solr

Objective:

Understand the functionalities of Apache Solr, including the creation of collections, data ingestion, and querying.

1. Environment Initialization

- Navigate to the Solr directory:
`cd solr`
- Start the Docker containers:
`docker-compose up -d`
- Access the Solr container:
`docker exec -it solr_solr_1 bash`
- If you can't access the Solr container, it could be due to a container name change. In this case use:
`docker exec -it solr-solr-1 bash`

2. Creating a Solr Collection

Exercise 1: Create a Solr collection named 'mynewcollection'.

```
/opt/solr/bin/solr create -c mynewcollection
```

Exercise 2: Verify that 'mynewcollection' has been successfully created.

- Access Solr's Web Interface at <http://localhost:8983/solr/> and check for 'mynewcollection' under the "Core Selector" dropdown.

Deliverable: Screenshot showing 'mynewcollection' in the Solr Web Interface.

3. Generating and Adding Data to the Collection

For this exercise, create a json file called products.json using this commands:

```
echo '[  
  {"id":"1", "name":"Product A", "category":"Electronics", "price":"100"},  
  {"id":"2", "name":"Product B", "category":"Books", "price":"20"}  
>' > /tmp/products.json
```

Exercise 3: Let's add the generated data to our collection.

```
/opt/solr/bin/post -c mynewcollection /tmp/products.json
```

Deliverable: Screenshot showing successful data ingestion messages in the terminal.

4. Querying the Collection

Exercise 4: Use Solr's query interface to retrieve all documents from 'mynewcollection'.

```
curl "http://localhost:8983/solr/mynewcollection/select?q=*:*"
```

Exercise 5: Query for products in the 'Electronics' category.

```
curl "http://localhost:8983/solr/mynewcollection/select?q=category:Electronics"
```

Deliverable: Screenshots of the JSON response for both queries.

5. Advanced Querying in Solr Web Interface

Access the Solr Web Interface for querying:

Query 'mynewcollection' in Solr Web Interface

Follow the link, and you'll be presented with a user-friendly interface to craft and execute your queries. Experiment with different parameters and filters.

Exercise 6: Try a faceted search to count the number of products in each category.

Exercise 7: Use the Web Interface to experiment further with various query parameters and filters.

Deliverable: Screenshots of different queries and their results from the Solr Web Interface.

6. Custom JSON Data

- Create a new collection in Solr
- Either generate JSON data or use JSON data from the web and add it to your collection.
- Query your collection.

Deliverable: Screenshots of different queries on your Solr collection and the query results from either the command line or the Solr Web Interface.

Shutting Down

Ensure all Docker containers are turned off with `docker-compose down` for each directory. If you're using google cloud, please shut down your virtual machine to preserve cloud costs.