Elasticsearch Advanced Tips

This is Heavy tech post, I will talk about most common tips of elasticsearch , putting together different points which are very important when you are using elastic search for your enterprise projects.

## Disable Doc\_Values when you don’t need them

Doc\_Values is column-oriented store in elastic search and it is enabled by default, created at index-time when a field is indexed, Elasticsearch adds the tokens to the inverted index for search.In another word contains same value as the \_source [Original Json Document] with different structure [column-oriented].

Doc\_values are used in several places such as Sorting, Aggregation and script which Column-oriented is way more efficient.

[\*] If you don’t need to sort or aggregation, or access the field value from a script, you HAVE to disable doc values in order to save disk space.

How to:

PUT demo\_index

{

"mappings": {

"\_doc": {

"properties": {

"username": {

"type": "keyword"

},

"session\_id ": {

"type": "keyword",

"doc\_values": false

}

}

}

}

}

## Synonymous

The concept of Synonymous is very interesting, handy and enriches the document in order to be retrieved with different combination of terms.

For example:

USA -> U.S.A

Sr. -> Senior

Dev -> Developer, Software Developer

…………………………………………………………………………..

Now Let’s imagine we would like to search “USA” and get all documents contains either U.S.A or U.S.A or USA, in order to solve this problem we have few options

1. At search time; we convert a query for USA into a query for USA OR “U.S.A”.
2. Using the synonym token filter in index time, we replace USA with the two terms usa and “U.S.A, then at search time we can find those terms.

Both solutions are valid but the first will drive us to have multiple responsibilities at application level and in my opinion it is not good practice.

With the second option we have centralized search engine will manage everything for us, converting the terms and make it available for search ☺

How to

PUT /vergnaty\_index

{

"settings": {

"analysis": {

"filter": {

"special\_synonyms": {

"type": "synonym",

"synonyms": [

"USA,U.S.A",

"Sr.,Senior"

]

}

},

"analyzer": {

"my\_synonyms": {

"tokenizer": "standard",

"filter": [“lowercase”, "special\_synonyms"]

}

}

}

}

}

GET /vergnaty\_index/\_analyze

{

"analyzer" : "my\_synonyms",

"text" : "Sr. Software developer in USA"

}

## Using Filtered Aliases with Routing

Aliases with filters provide an easy way to create different "views" of the same index, all documents

Are routed to a particular shard in an index and by default is used the document’s \_id.

Having Filtered Aliases combined with \_routing can reduce the impact of searches. Instead of having to fan out a search request to all the shards in an index, the request can be sent to just the shard that matches the specific routing values.