

| Equivalent Terms |               |
|------------------|---------------|
| SQL              | ElasticSearch |
| Database         | Index         |
| Table            | Type          |
| Row              | Document      |
| Column           | Field         |

| Definitions (?) |   |
|-----------------|---|
| Index           | Collection of one or more shards                                      |
| Shard           | Apache Lucene Instance  |
| Cluster         | One or more nodes   |
| Nodes           | Server instances  |
| Routing         | A way to control what shard a document goes to (used for performance) |
| Segment         | Internal lucene data block  |

| Resilience & Scalability |   |
|--------------------------|---|
| Shard(s)                 | One Primary, N Replicas                                     |
| Index                    | Made of one or more shards                                  |
| Cluster                  | Made of one or more nodes                                   |
| Nodes                    | A single server in a cluster                                |
| Shard                    | Logical partition of an index across the nodes in a cluster |
| Replica                  | Number of copies of the primary shards in a cluster         |
| Total shards in cluster  | = primary shards * (replica count + 1)                      |
| Shards per node          | = Total shards in cluster / # of nodes                      |

| Things Handled By  |                     |
|--------------------|---------------------|
| ElasticSearch      | Apache Lucene       |
| Node management    | Searching           |
| Shard management   | Indexing            |
| Health monitoring  | Stemming            |
| Parallelizing Work | Serialized Workload |

| Parent-Child vs Nested |                                   |
|------------------------|-----------------------------------|
| Sharding               | Children reside in Parent's shard |
| Parent-Child           | Cheap to update, slow to query    |
| Nested                 | Cheap to query, slow to update    |

| Crud Commands                                 |   |
|---|---|
| PUT<br>server/index/type/ID<br>JSON           | Creates doc from JSON using ID              |
| POST<br>server/index/type<br>JSON             | Creates doc from JSON and generates ID      |
| POST<br>server/index/type/ID/_<br>update JSON | Updates part of the document matching on ID |
| GET<br>server/index/type/ID                   | Retrieves doc matching ID                   |
| DELETE<br>server/index/type/ID                | Deletes doc matching ID                     |

| Search Commands                            |         |                                      |
|--|---------|--------------------------------------|
| GET<br>server/index/type/_s<br>earch?q=XXX | Simple  | Search index\type with query XXX     |
| POST<br>server/index/type/_s<br>earch JSON | Complex | Search index\type with query in JSON |

| Big Ideas      |  |
|----------------|--|
| Forward Index  | Maps a document to a set of words                                  |
| Inverted Index | Maps a word to a set of documents                                  |
| Stemming       | Converting a word to it's root, ie raining -> rain                 |
| Synonyms       | Converts a word to an appropriate equivalent, ie Sunday -> weekend |

**Lucene uses an *Inverted Index* to find documents so quickly.**

**Lucene uses *stems* for it's indexes and search terms to increase hit ratio for searches.**

**Lucene uses *synonyms* to increase hit ratio for searches.**



By **Louis Cribbins**  
[cheatography.com/louis-cribbins/](https://cheatography.com/louis-cribbins/)

Not published yet.  
 Last updated 10th April, 2018.  
 Page 1 of 1.

Sponsored by **Readability-Score.com**  
 Measure your website readability!  
<https://readability-score.com>