

## ElasticSearch Interview Questions

Posted on [February 17, 2018](#) by [ProTechSkills](#)



### Categories

[.Net](#) (2)  
[Big Data](#) (40)  
[Elastic Search](#) (1)  
[Hadoop](#) (27)  
[Hadoop Administration](#) (20)  
[Hadoop 1](#) (6)  
[Hadoop 2](#) (11)  
[HDFS](#) (8)  
[MapReduce](#) (6)  
[YARN](#) (5)  
[Hadoop Ecosystem](#) (15)

### What is ElasticSearch?

Elasticsearch is a search engine based on Lucene. It has a distributed, multitenant-able full-text search engine. Elasticsearch is developed in Java and is released as open source under the terms of the License.

### What is an index in ElasticSearch?

An index is a collection of documents that have somewhat similar characteristics. For example, you can have an index for customer data, another index for a product catalog, and yet another index for logs. An index is identified by a name (that must be all lowercase) and this name is used to refer to the index when performing indexing, update, and delete operations against the documents in it.

In a single cluster, you can define as many indexes as you want.

### What is Type (Mapping Type) in Index of ElasticSearch?

×

### Demo Class

Please share your details to attend a demo class

\* Name

\* Email

\* Phone

\* Course

Register

Connecto

A type used to be a logical category/partition of your index to allow you to store different types of documents in the same index, e.g. one type for users, another type for blog posts.

It has been marked as deprecated and will no longer be possible to create multiple types in an index, and the whole concept of types will be removed in a Elasticsearch 7.x version.

### What is the use of field level attributes- index and store?

The index is employed for searching. Indexed fields are transformed during analysis, and cannot retrieve the original data when necessary.

Store implies the data stored by Lucene, which will again return when necessary. Stored fields are not searchable.

### What is an Analyzer in Elasticsearch?

While indexing data, it is transformed internally via the defined Analyzer for the index.

Analyzers are made of one Tokenizer, preceded by CharFilters and zero or many TokenFilters. On the other hand, analysis module refers Analyzers under the name of mapping definitions or any APIs.

Elasticsearch is prebuilt with analyzers that are ready to use. However, you can integrate the built in character, token filters, along with custom filters to create custom analyzers.

### What is Character Filter in Elasticsearch Analyzer?

A character filter obtains the ideal text as stream of characters and modifies it by adding, deleting, or altering characters. For example, any character filter in usage has the ability to convert Hindu-Arabic numerals (०१२३४५६७८९) into Arabic-Latin numerals (0123456789) or even sometimes strip HTML elements via the stream.

### What is Token filters in Elasticsearch Analyzer?

A token filter obtains the token stream, later on add, delete, or modify the tokens. For instance, a lowercase token filter modifies all

[QTP \(4\)](#)

[Videos \(1\)](#)

### Recent Posts

[ElasticSearch Interview Questions](#)

[Apache Pig Data Types](#)

[Getting started with MapReduce](#)

[Programming](#)

[Import Incremental](#)

[Data using Sqoop](#)

[Writing Custom](#)

[Combiner in](#)

[MapReduce](#)

### Archives

[February 2018](#)

[November 2016](#)

[August 2016](#)

[July 2016](#)

[June 2016](#)

[May 2016](#)

[February 2016](#)

[October 2015](#)

[June 2015](#)

×

## Demo Class

Please share your details to attend a demo class

\* Name

\* Email

\* Phone

\* Course

Register

Connecto

lowercase, a stop token filter deletes stop words, and a synonym token filter includes synonyms into the token stream.

Token filters will be unable to change the position or character offsets of any certain token.

## What is a Tokenizer?

Tokenizers break down a string into stream of tokens. A single tokenizer split the string into terms when working with punctuation and whitespace. Elasticsearch has a number of built in tokenizers which can be used to build custom analyzers.

## using Selenium WebDriver?

Phoenix Locksmiths on  
HDFS High Availability  
Configuration

Phoenix Locksmith on  
Testing Principles

## Meta

Log in  
Entries RSS  
Comments RSS  
WordPress.org

## What are the advantages of Elasticsearch?

- Elasticsearch is compatible on any platform.
- Elasticsearch is Near Real Time (NRT), making it searchable on engine.
- Elasticsearch cluster is distributed, scalable and easy to integrate.
- Elasticsearch REST uses JSON objects, making it to invoke the Elasticsearch server along with different programming languages.
- Elasticsearch supports every document type except text rendering.

## What is Elasticsearch REST API and use of it?


Elasticsearch provides a very comprehensive and powerful REST API that you can use to interact with your cluster. Among the few things that can be done with the API are as follows:

Check your cluster, node, and index health, status, and statistics

## Administer your cluster, node, and index data and metadata

Perform CRUD (Create, Read, Update, and Delete) and search against your indexes

Execute advanced search operations viz. aggregations, filtering, scripting, sorting, among many others



## Demo Class

Please share your details to attend a demo class

\*

Name

\*


Email

\*

Phone


\*

Course



Register

Connecto



## Does Elasticsearch have a schema?

Yes, Elasticsearch can have a schema. A schema is a description of one or more fields that describes the document type and how to handle the different fields of a document. The schema in Elasticsearch is a mapping that emphasizes the JSON document fields and other data type, as well as Lucene indexes under the hood. Because of this, in Elasticsearch terms, we usually call this schema a “mapping”.

## What is a cluster in Elasticsearch?

Cluster is a collection of nodes that holds data together and enables indexing and search abilities across each. Each cluster is recognized by a unique default name i.e. “elasticsearch”. This name is important because a node can only be part of a cluster if the node is set up to join the cluster by its name.

## What is a node in Elasticsearch?

Node is a minute server and forms a part of the cluster. It stores the data and enjoys the clusters indexing and search functionalities.

## What is Ingest Node in Elasticsearch?

Ingest nodes can execute pre-processing an ingest pipeline. It effectively transform and works on the document prior to indexing. Dedicated ingest nodes mark the master and data nodes either `node.ingest=true` or `node.data=true`.

## What is Elasticsearch Data Node?

Data nodes hold shards that handle indexed documents. They execute data related CRUD and search aggregation operations. To make a node a Data Node, set `node.data=true`.

Data Node operations are I/O-, memory-, and CPU-intensive. They benefit the separation of the master and data roles.

## What is Master Node and Master Eligible Node in Elasticsearch?

×

### Demo Class

Please share your details to attend a demo class


\* Name

\* Email

\* Phone

\* Course ▾

Register

Connecto 

Master Node control cluster wide operations like to create or remove an index, track nodes of cluster, and decide to allocate shards on nodes. It is important for cluster health to have a stable master node. Master Node elected based on configuration properties `node.master=true` (Default).

Master Eligible Node decide based on below configuration

`discovery.zen.minimum_master_node` : number (default 1)

and above number is decided based on  $(\text{master\_eligible\_nodes} / 2) + 1$

## What is Tribe Node and Coordinating Node in Elasticsearch?

Tribe node connect variant clusters and execute search operations across each connected clusters. This node is configured by settings *tribe*.

Coordinating Node is just like a Smart Load balancer that handle master duties, to hold data, and pre-process documents, then you are left with a coordinating node that can only route requests, handle the search reduce phase, and distribute bulk indexing.

Every node can be termed as a coordinating node which has all three `node.data`, `node.ingest` and `node.master`, set to false. This node is impossible to disable as it possess enough memory and CPU to deal with the gather phase.

## What is inverted index in Elasticsearch?

Inverted Index is backbone of Elasticsearch which make full-text fast. Inverted index consists of a list of all unique words that documents and for each word, maintain a list of documents positions in which it appears.

For Example: There are two documents and having content a

1: FacingIssuesOnIT is for ELK.

2: If ELK check FacingIssuesOnIT.

×

## Demo Class

Please share your details to attend a demo class


\* Name

\* Email

\* Phone

\* Course ▾

Register

Connecto 

To make inverted index each document will split in words (also called as terms or token) and create below sorted index .

Term	Doc_1	Doc_2
FacingIssuesOnIT	X	X
Is	X	
For	X	
ELK	X	X
If		X
Check		X

Now when we do some full-text search for String will sort documents based on existence and occurrence of matching counts.

Usually in Books we have inverted indexes on last pages. Based on the word we can thus find the page on which the word exists.

## What is a shard?

An index can potentially store a large amount of data that can exceed the hardware limits of a single node. For example, a single index of a billion documents taking up 1TB of disk space may not fit on the disk of a single node or may be too slow to serve search requests from a single node alone.

To solve this problem, Elasticsearch provides the ability to subdivide an index into multiple pieces called shards. When you create an index, you can simply define the number of shards that you want. Each shard is itself a fully-functional and independent “index” that can be hosted on a node in the cluster.

Sharding is important for two primary reasons:

- It allows you to horizontally split/scale your content vertically
- It allows you to distribute and parallelize operations across multiple nodes (potentially on multiple nodes) thus increasing performance/throughput

## What is a replica?

×

## Demo Class

Please share your details to attend a demo class

\* Name


\* Email

\* Phone

\* Course ▼

Register

Connecto



Elasticsearch allows you to make one or more copies of your index's shards into what are called replica shards, or replicas for short.

Replication is important for two primary reasons:

- It provides high availability in case a shard/node fails. For this reason, it is important to note that a replica shard is never allocated on the same node as the original/primary shard that it was copied from.
- It allows you to scale out your search volume/throughput since searches can be executed on all replicas in parallel.

## What is a document in ElasticSearch?

A document is a basic unit of information that can be indexed. For example, you can have a document for a single customer, another document for a single product, and yet another for a single order. This document is expressed in JSON (JavaScript Object Notation) which is a ubiquitous internet data interchange format.

Within an index/type, you can store as many documents as you want. Note that although a document physically resides in an index, a document actually must be indexed/assigned to a type inside an index.

## What are the basic operations you can perform on a document?

The following operations can be performed on documents

INDEXING A DOCUMENT USING ELASTICSEARCH.

FETCHING DOCUMENTS USING ELASTICSEARCH.

UPDATING DOCUMENTS USING ELASTICSEARCH.

DELETING DOCUMENTS USING ELASTICSEARCH.

---

Share this:

×

## Demo Class

Please share your details to attend a demo class

\* Name


\* Email

\* Phone

\* Course

▼

Register

Connecto

This entry was posted in [Big Data](#), [Elastic Search](#) and tagged [elasticsearch](#), [elasticsearch interview questions](#), [elasticsearch interview questions and answers for experienced](#), [elasticsearch questions and answers](#), [elk stack interview questions](#). Bookmark the [permalink](#).

← [Apache Pig Data Types](#)

## Leave a Reply

Your email address will not be published. Required fields are marked \*

COMMENT

NAME \*

EMAIL \*

WEBSITE

> [pe to Unlock](#)

Post Comment

Get in touch with us today! Attend a FR Class.

Address

Phone

+91 84476 03880

Email

[training@](#)

### Demo Class

Please share your details to attend a demo class

\* Name

\* Email

\* Phone

\* Course

Register

Connecto