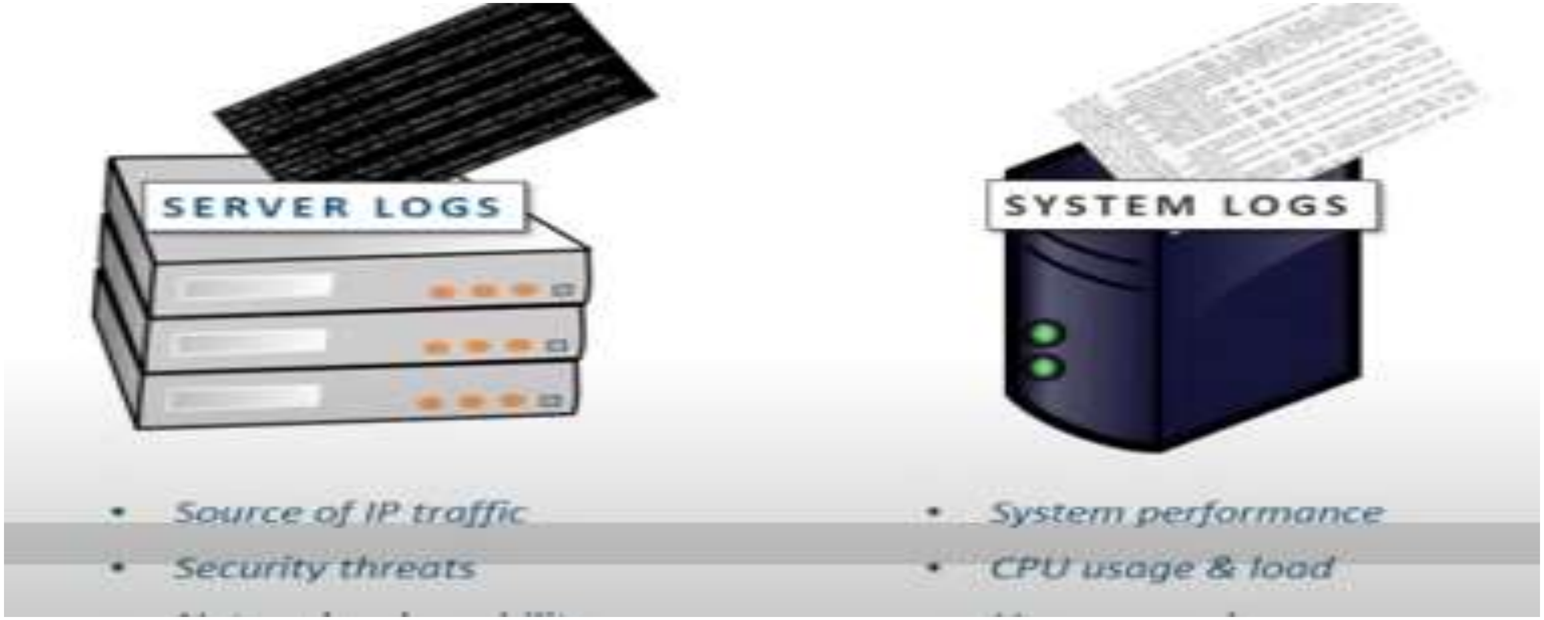
Splunk

Ravindra Kudache



SERVER LOGS

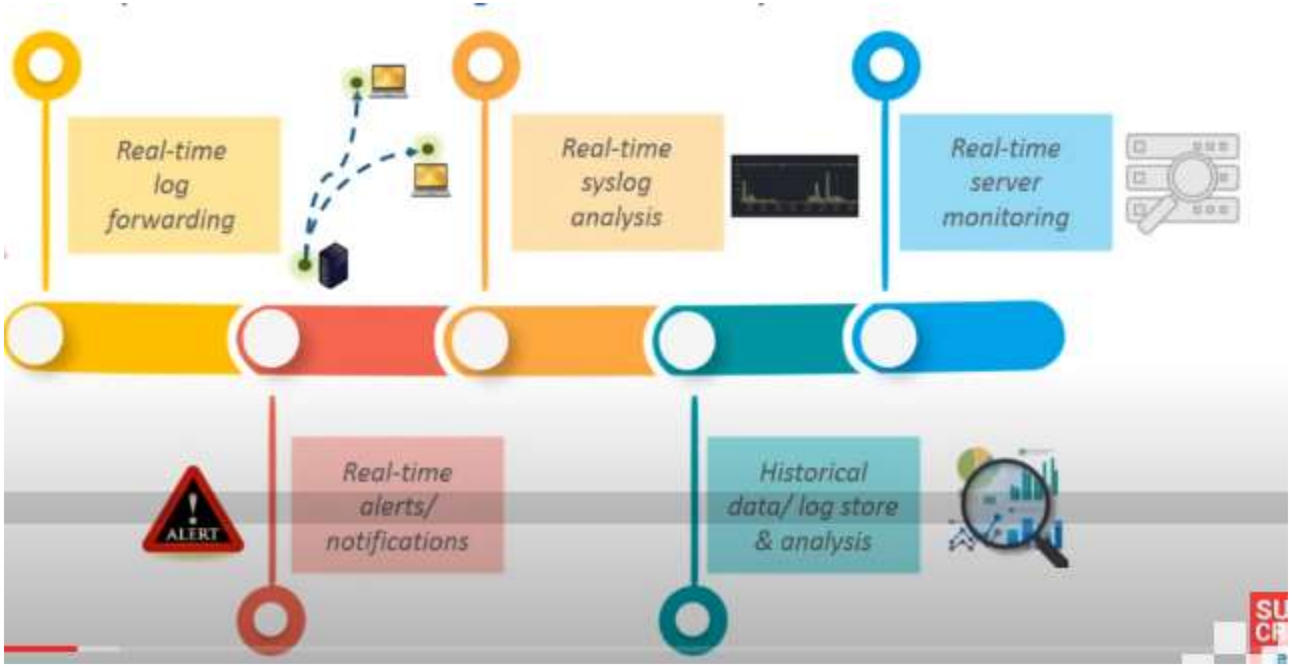
- *Source of IP traffic*
- *Security threats*

SYSTEM LOGS

- *System performance*
- *CPU usage & load*

```
61.245.165.56 - - [24/May/2012:10:07:26 -0400] "GET /images/web/2009/banner.png HTTP/1.1" 200 523
Firefox/13.0"
178.77.103.6 - - [24/May/2012:10:08:02 -0400] "GET /blog HTTP/1.1" 200 36520 "-" "Tiny Tiny RSS/1
66.177.138.85 - - [24/May/2012:10:08:16 -0400] "GET /blog/geekery/comcast-dns-hijack-breaks-thing
sa=t&rct=j&q=dns%20hikack%20test&source=web&cd=1&ved=0CFwQFjAA&url=http%3A%2F%2Fwww.semicomplete.
T7H_DoKogwfvpaCj0Q&usg=AFQjCNFyb7tcVQ7gfEoM4uwEcmuAfA_R7g&cad=rja" "Mozilla/5.0 (Windows
66.177.138.85 - - [24/May/2012:10:08:56 -0400] "GET /reset.css HTTP/1.1" 200 1015 "htt
rv:12.0) Gecko/20100101 Firefox/12.0"
66.177.138.85 - - [24/May/2012:10:08:56 -0400] "GET /images/jordan-80.png HTTP/1.1"
WOW64; rv:12.0) Gecko/20100101 Firefox/12.0"
66.177.138.85 - - [24/May/2012:10:08:56 -0400] "GET /style2.css HTTP/1.1" 200 4877
rv:12.0) Gecko/20100101 Firefox/12.0"
66.177.138.85 - - [24/May/2012:10:09:05 -0400] "GET /favicon.ico HTTP/1.1" 200 3638
66.177.138.85 - - [24/May/2012:10:09:17 -0400] "GET /images/web/2009/banner.png HTT
Firefox/12.0"
83.173.180.145 - - [24/May/2012:10:09:46 -0400] "GET /articles/dynamic-dns-with-d
83.173.180.145 - - [24/May/2012:10:09:46 -0400] "GET /reset.css HTTP/1.1" 200 101
Gecko/20100101 Firefox/12.0"
83.173.180.145 - - [24/May/2012:10:09:46 -0400] "GET /style2.css HTTP/1.1" 200 4877 "
Gecko/20100101 Firefox/12.0"
83.173.180.145 - - [24/May/2012:10:09:46 -0400] "GET /images/jordan-80.png HTTP/1.1"
Gecko/20100101 Firefox/12.0"
83.173.180.145 - - [24/May/2012:10:09:47 -0400] "GET /favicon.ico HTTP/1.1" 200 3638 "
195.122.130.126 - - [24/May/2012:10:09:47 -0400] "GET /projects/xdotool HTTP/1.1" 301 3
83.173.180.145 - - [24/May/2012:10:09:47 -0400] "GET /images/web/2009/banner.png HTTP/1
Firefox/12.0"
195.122.130.126 - - [24/May/2012:10:09:47 -0400] "GET /projects/xdotool/ HTTP/1.1" 251
```

But 1000s of such logs are generated every minute
Check security threats in real-time!!!
Analyze business metrics in real-time!!!



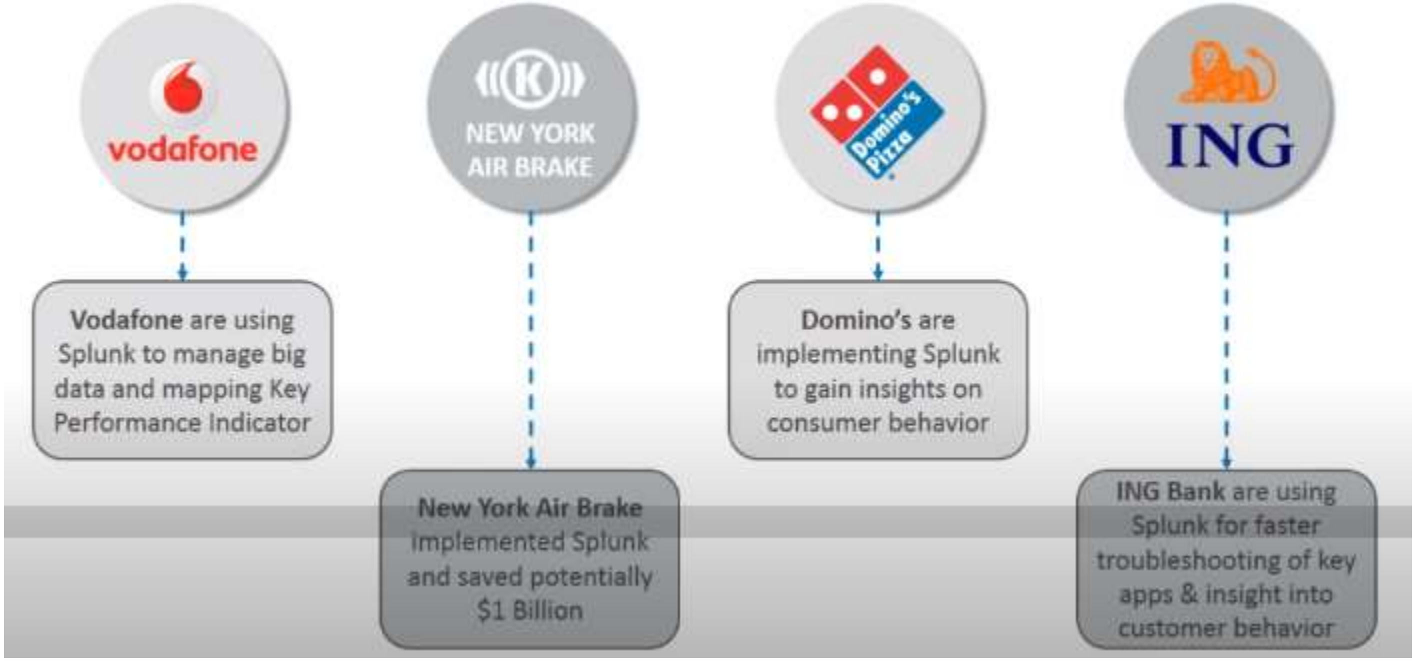
Real-time log forwarding

Real-time alerts/ notifications

Real-time syslog analysis

Historical data/ log store & analysis

Real-time server monitoring

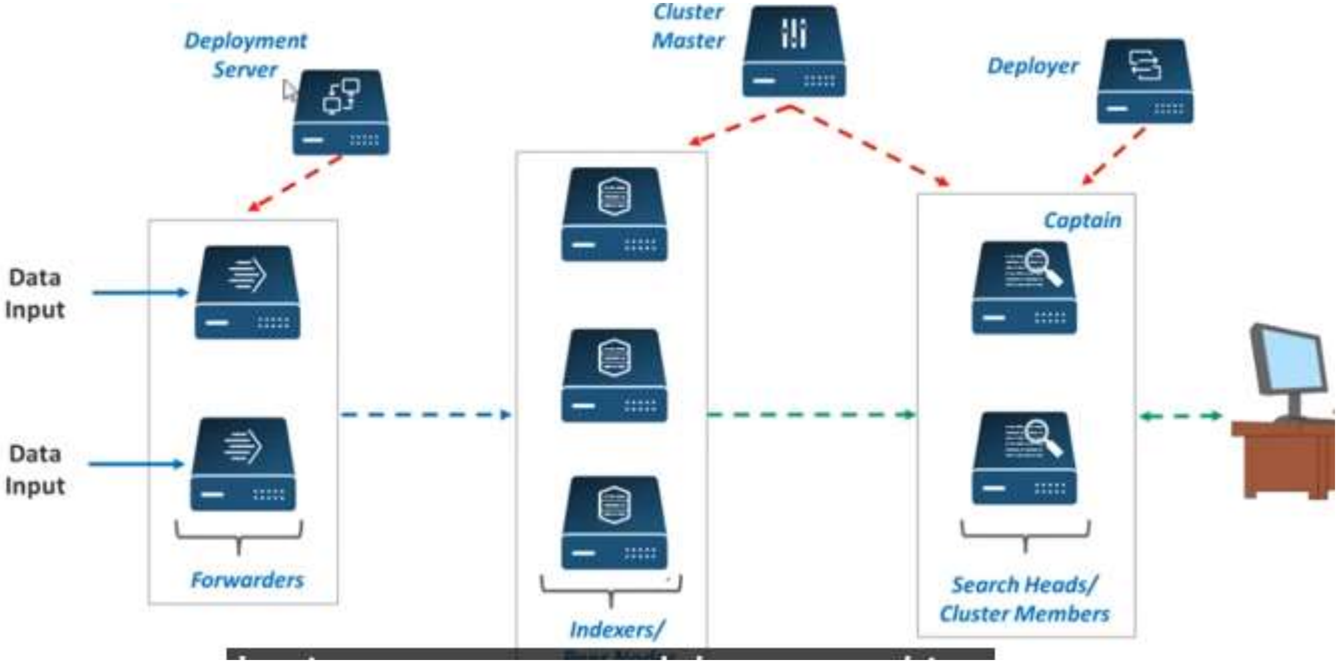


Vodafone are using Splunk to manage big data and mapping Key Performance Indicator

New York Air Brake implemented Splunk and saved potentially $1 Billion

Domino's are implementing Splunk to gain insights on consumer behavior

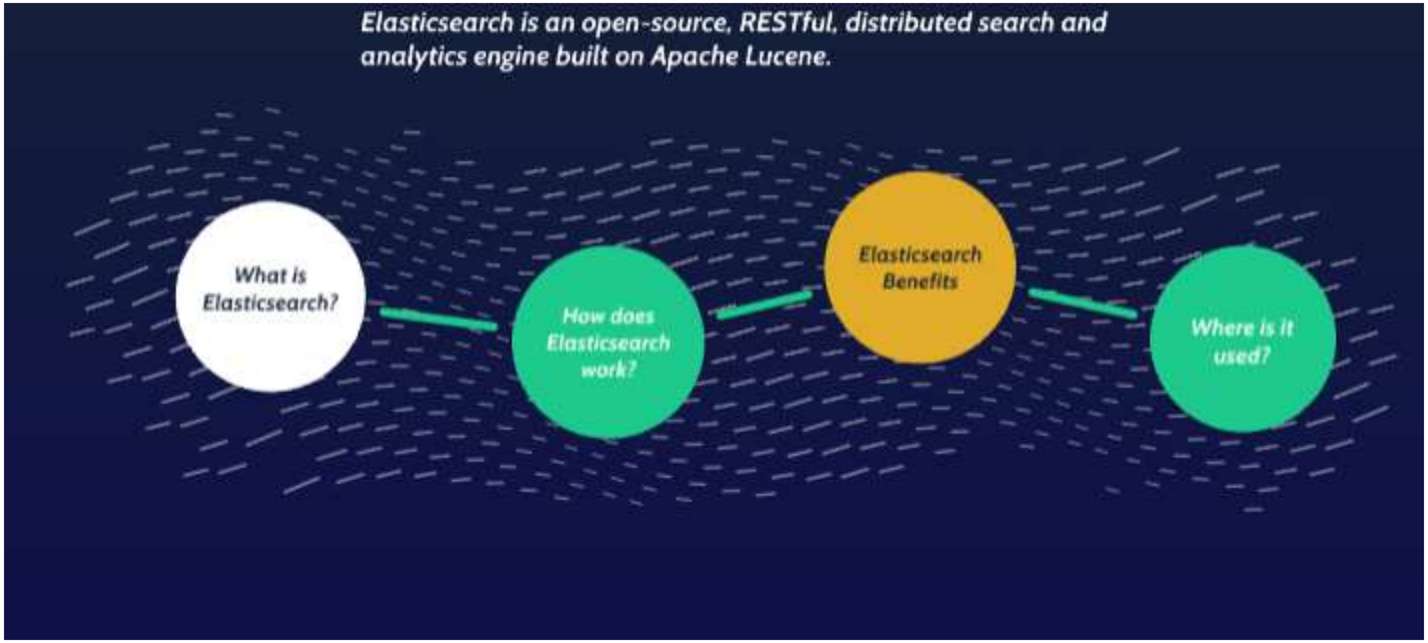
ING Bank are using Splunk for faster troubleshooting of key apps & insight into customer behavior


Deployment Server
Cluster Master
Deployer
Captain
Data Input
Data Input
Forwarders
Indexers/
Search Heads/ Cluster Members

Install Splunk

App Dynamic

APM211+-+Introduction+to+AppDynamics+-+Student+Guide.pdf



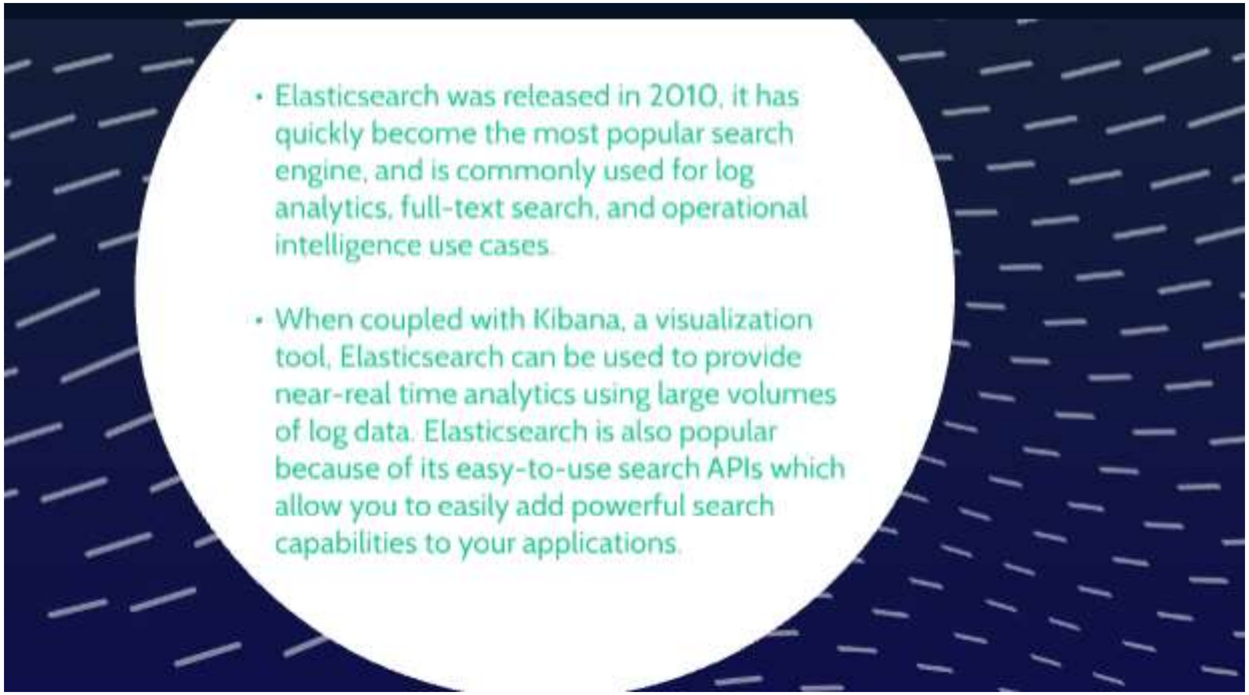
Elasticsearch is an open-source, RESTful, distributed search and analytics engine built on Apache Lucene.

What is Elasticsearch?

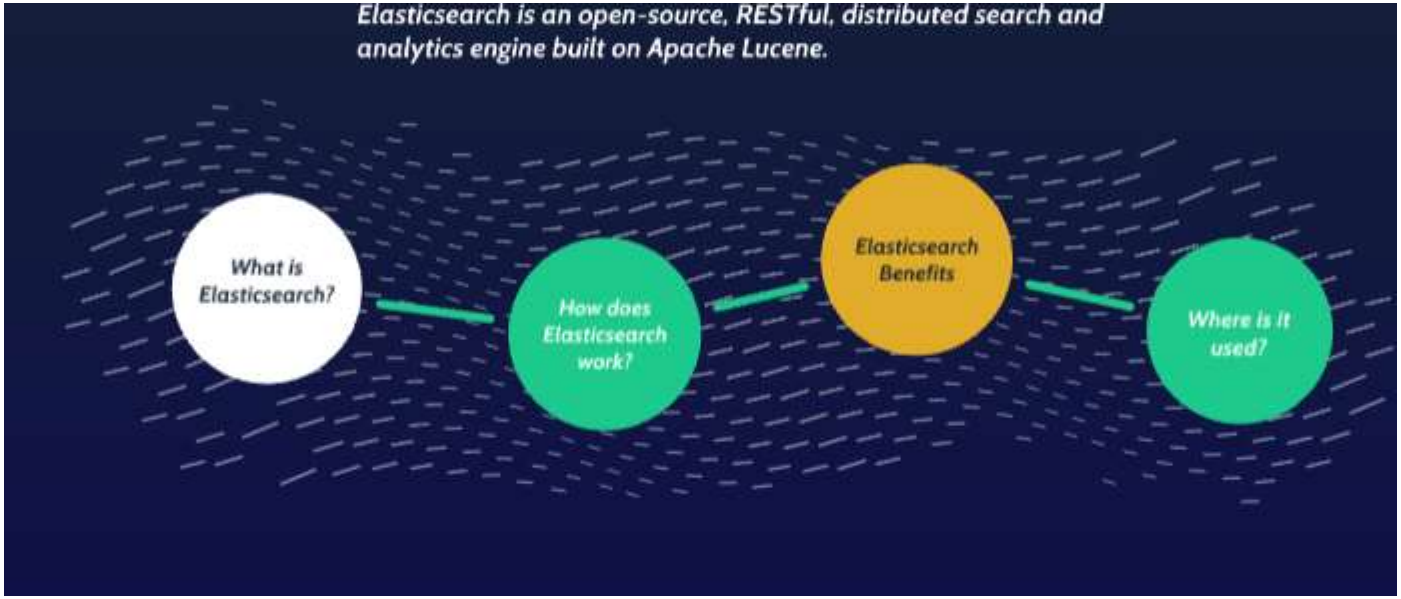
How does Elasticsearch work?

Elasticsearch Benefits

Where is it used?



- Elasticsearch was released in 2010, it has quickly become the most popular search engine, and is commonly used for log analytics, full-text search, and operational intelligence use cases.
- When coupled with Kibana, a visualization tool, Elasticsearch can be used to provide near-real time analytics using large volumes of log data. Elasticsearch is also popular because of its easy-to-use search APIs which allow you to easily add powerful search capabilities to your applications.



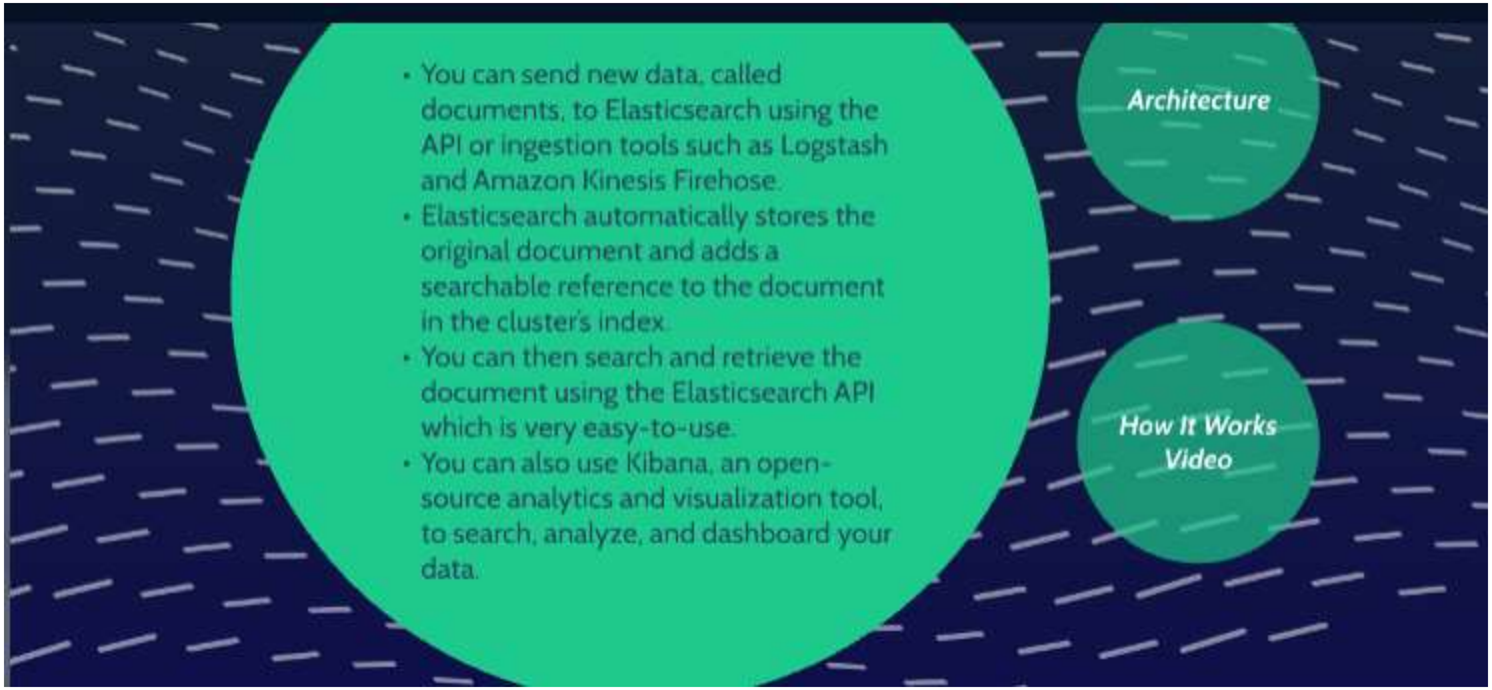
Elasticsearch is an open-source, RESTful, distributed search and analytics engine built on Apache Lucene.

What is Elasticsearch?

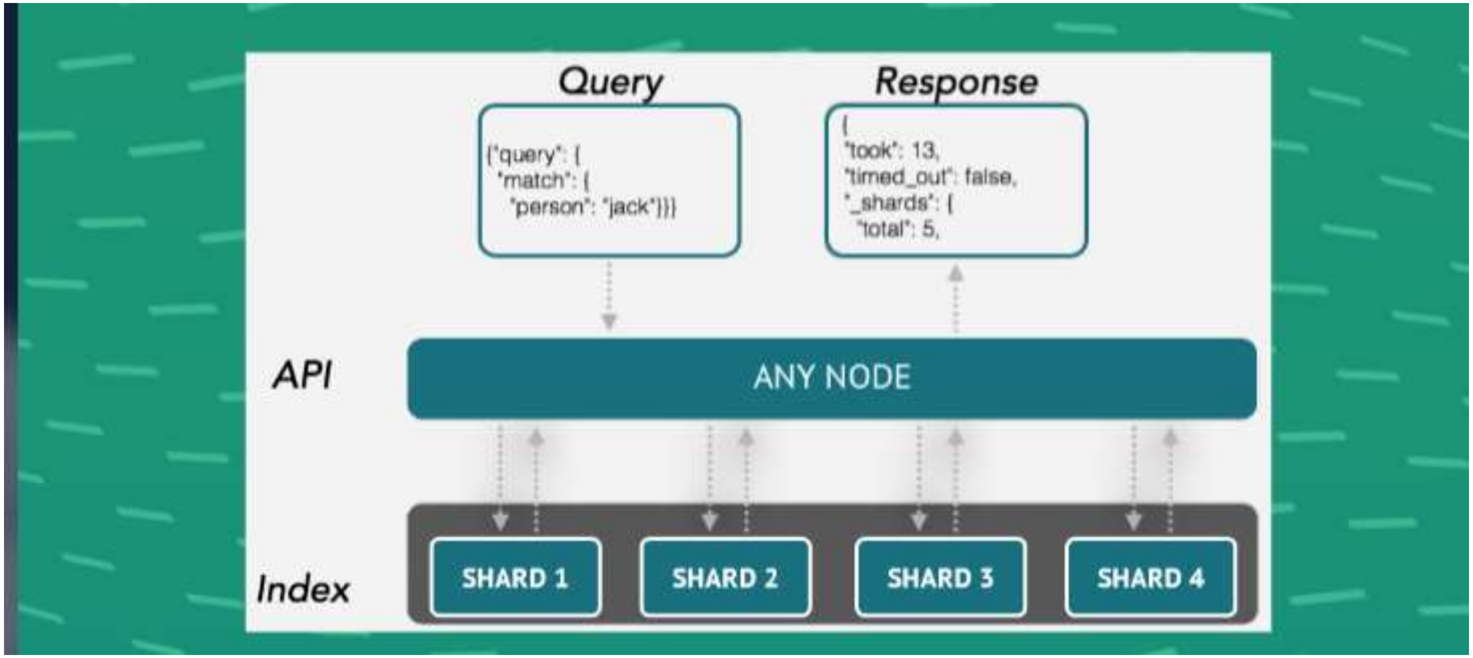
How does Elasticsearch work?

Elasticsearch Benefits

Where is it used?



- You can send new data, called documents, to Elasticsearch using the API or ingestion tools such as Logstash and Amazon Kinesis Firehose.
- Elasticsearch automatically stores the original document and adds a searchable reference to the document in the cluster's index.
- You can then search and retrieve the document using the Elasticsearch API which is very easy-to-use.
- You can also use Kibana, an open-source analytics and visualization tool, to search, analyze, and dashboard your data.



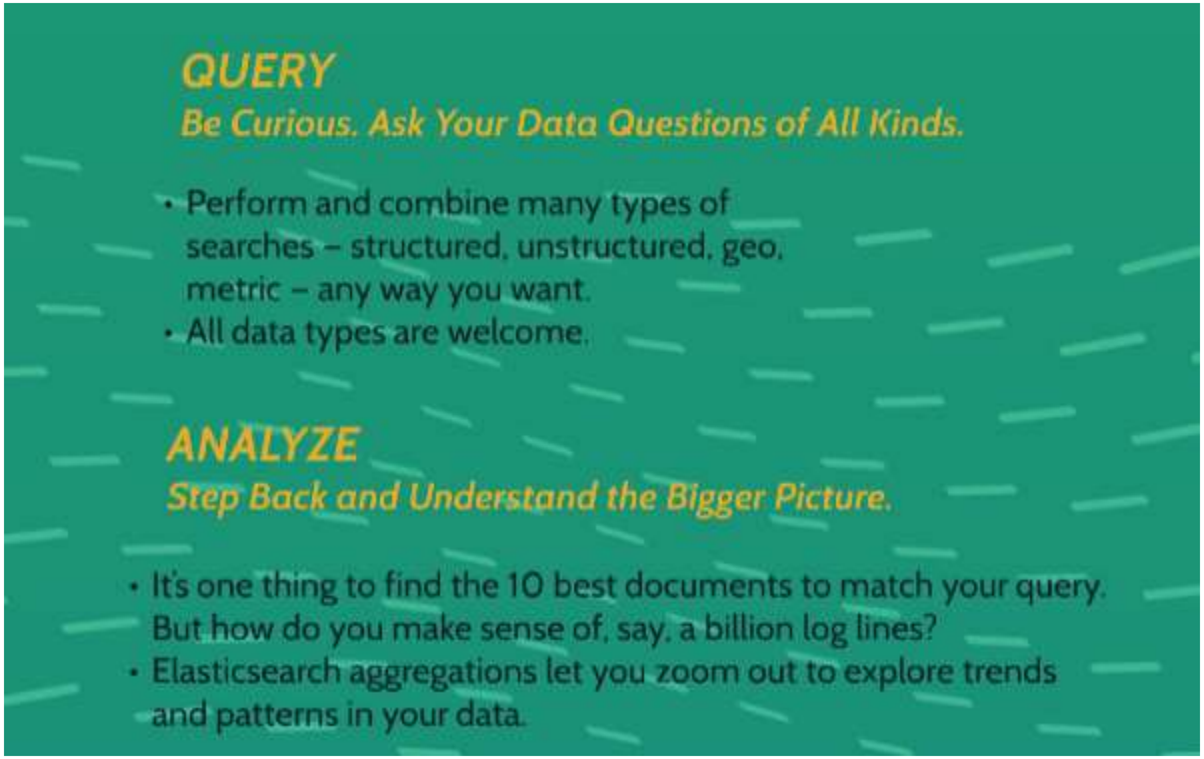
```
{"query": {
  "match": {
    "person": "jack"}}}
```

```
{
"took": 13,
"timed_out": false,
"_shards": {
  "total": 5,
```

BENEFITS

QUERY
ANALYZE
SPEED
SCALABILITY
COMPLIMENTARY TOOLING AND PLUG-INS
NEAR REAL-TIME INDEX UPDATES
CLIENT LIBRARIES
MACHINE LEARNING
SECURITY

Speed and Scalability
Complimentary tooling and plug-ins
Real-time updates and Client libs
Machine learning and Security



QUERY

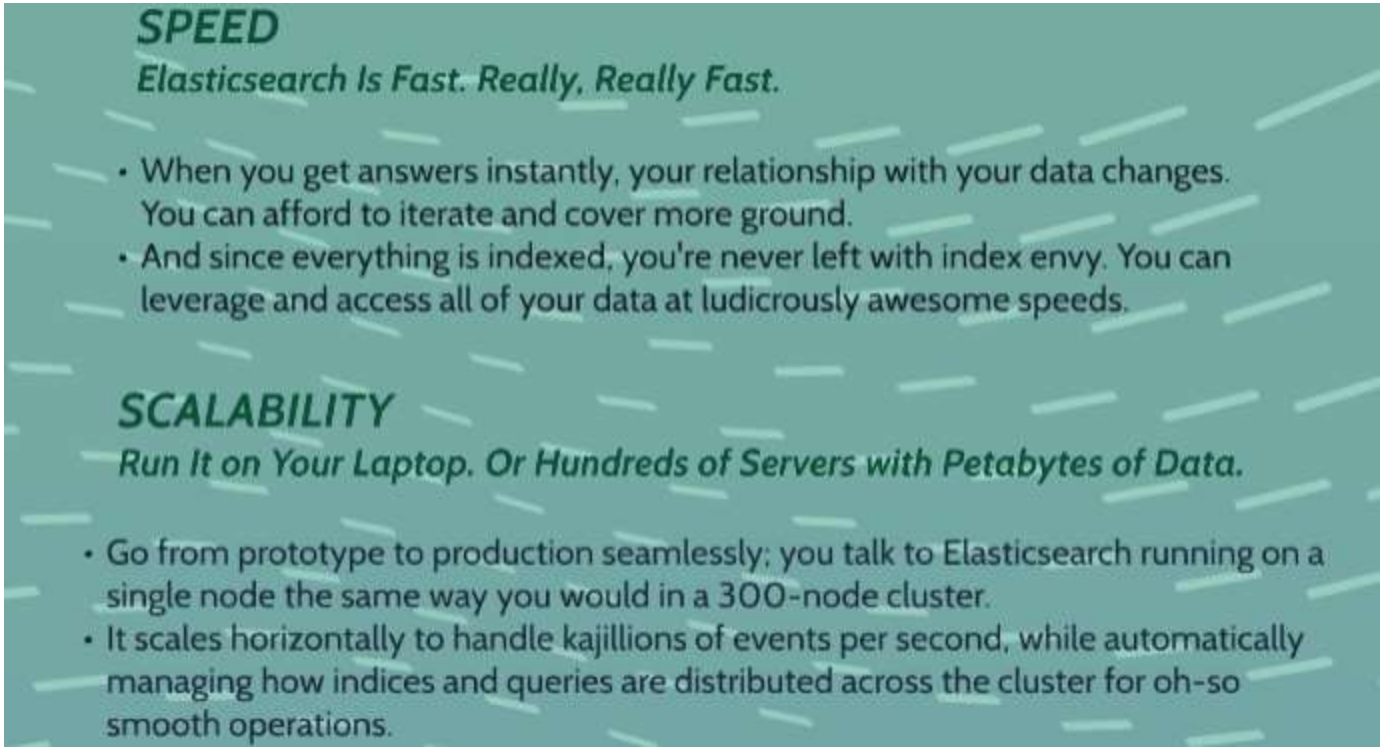
Be Curious. Ask Your Data Questions of All Kinds.

- Perform and combine many types of searches – structured, unstructured, geo, metric – any way you want.
- All data types are welcome.

ANALYZE

Step Back and Understand the Bigger Picture.

- It's one thing to find the 10 best documents to match your query. But how do you make sense of, say, a billion log lines?
- Elasticsearch aggregations let you zoom out to explore trends and patterns in your data.



SPEED

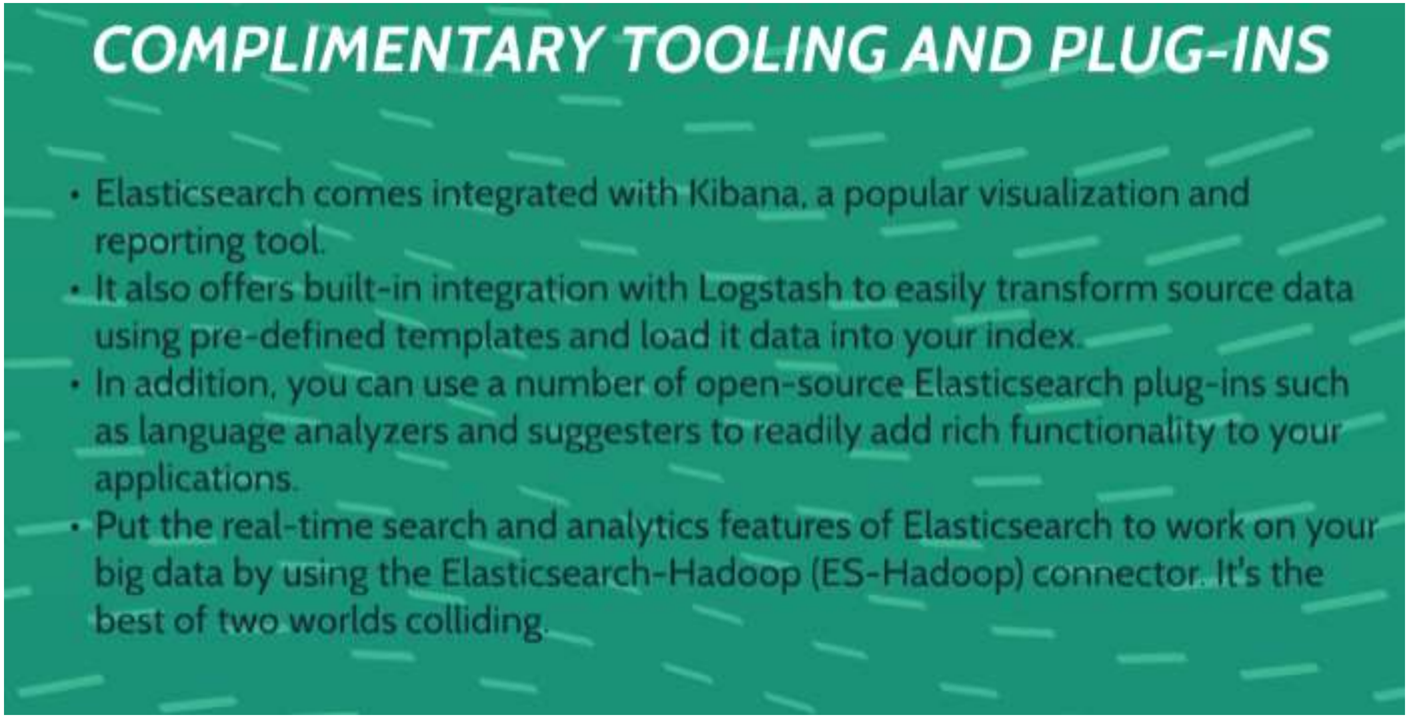
Elasticsearch Is Fast. Really, Really Fast.

- When you get answers instantly, your relationship with your data changes. You can afford to iterate and cover more ground.
- And since everything is indexed, you're never left with index envy. You can leverage and access all of your data at ludicrously awesome speeds.

SCALABILITY

Run It on Your Laptop. Or Hundreds of Servers with Petabytes of Data.

- Go from prototype to production seamlessly; you talk to Elasticsearch running on a single node the same way you would in a 300-node cluster.
- It scales horizontally to handle kajillions of events per second, while automatically managing how indices and queries are distributed across the cluster for oh-so smooth operations.



COMPLIMENTARY TOOLING AND PLUG-INS

- Elasticsearch comes integrated with Kibana, a popular visualization and reporting tool.
- It also offers built-in integration with Logstash to easily transform source data using pre-defined templates and load it data into your index.
- In addition, you can use a number of open-source Elasticsearch plug-ins such as language analyzers and suggesters to readily add rich functionality to your applications.
- Put the real-time search and analytics features of Elasticsearch to work on your big data by using the Elasticsearch-Hadoop (ES-Hadoop) connector. It's the best of two worlds colliding.

NEAR REAL-TIME INDEX UPDATES

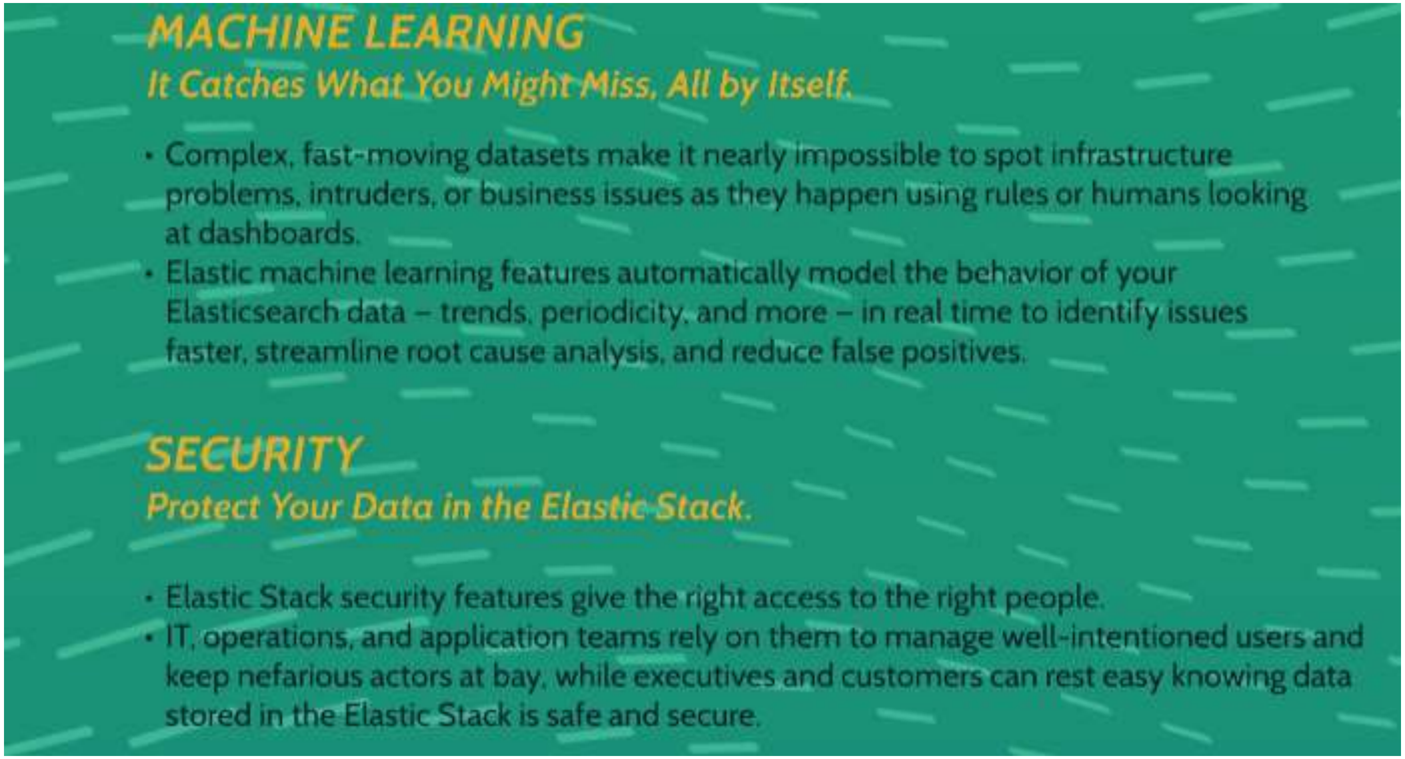
Elasticsearch Is Fast. Really, Really Fast.

- Elasticsearch index updates such as adding a new document to the index usually take one second or less before the updated data is available for search.
- This lets you use Elasticsearch for near real-time use cases such as application monitoring and anomaly detection.

CLIENT LIBRARIES

Support for your Favorite Development Language

- A variety of open source clients are available for Elasticsearch developers.
- Supported languages include Java, Python, .NET, SQL, PHP, JavaScript, Node.js, Ruby, and many others.



MACHINE LEARNING

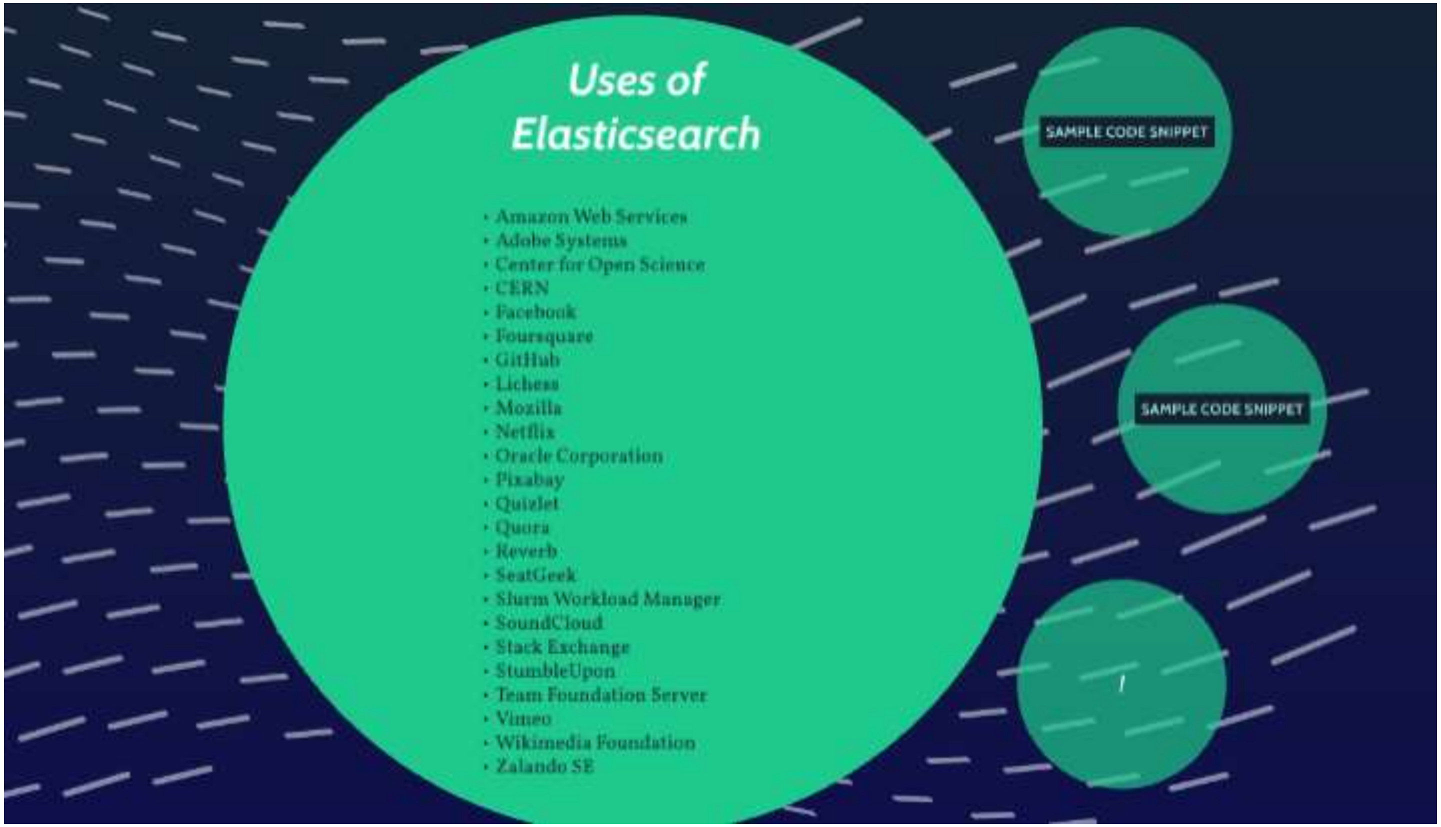
It Catches What You Might Miss, All by Itself.

- Complex, fast-moving datasets make it nearly impossible to spot infrastructure problems, intruders, or business issues as they happen using rules or humans looking at dashboards.
- Elastic machine learning features automatically model the behavior of your Elasticsearch data – trends, periodicity, and more – in real time to identify issues faster, streamline root cause analysis, and reduce false positives.

SECURITY

Protect Your Data in the Elastic Stack.

- Elastic Stack security features give the right access to the right people.
- IT, operations, and application teams rely on them to manage well-intentioned users and keep nefarious actors at bay, while executives and customers can rest easy knowing data stored in the Elastic Stack is safe and secure.

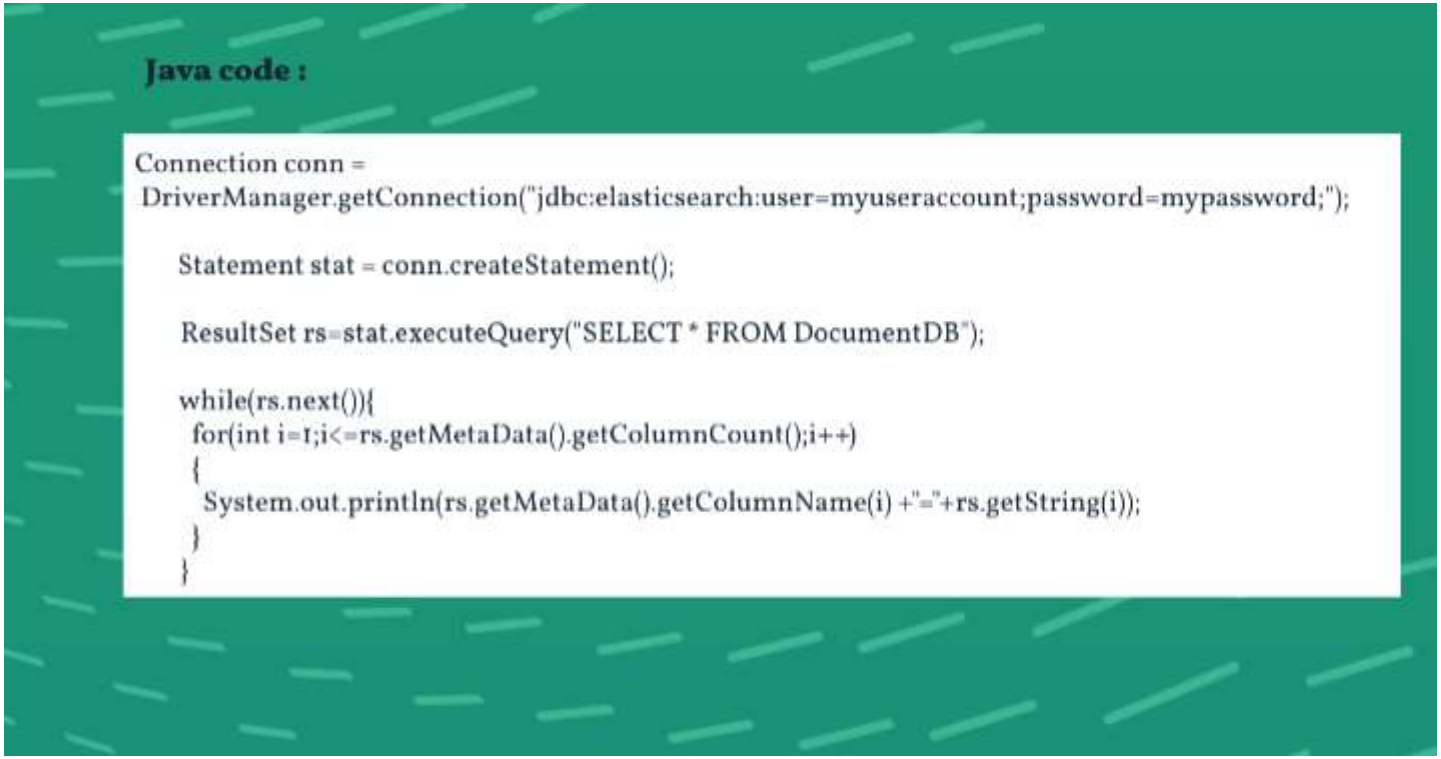


Uses of Elasticsearch

- Amazon Web Services
- Adobe Systems
- Center for Open Science
- CERN
- Facebook
- Foursquare
- GitHub
- Lichess
- Mozilla
- Netflix
- Oracle Corporation
- Pixabay
- Quizlet
- Quora
- Reverb
- SeatGeek
- Slurm Workload Manager
- SoundCloud
- Stack Exchange
- StumbleUpon
- Team Foundation Server
- Vimeo
- Wikimedia Foundation
- Zalando SE

SAMPLE CODE SNIPPET

SAMPLE CODE SNIPPET



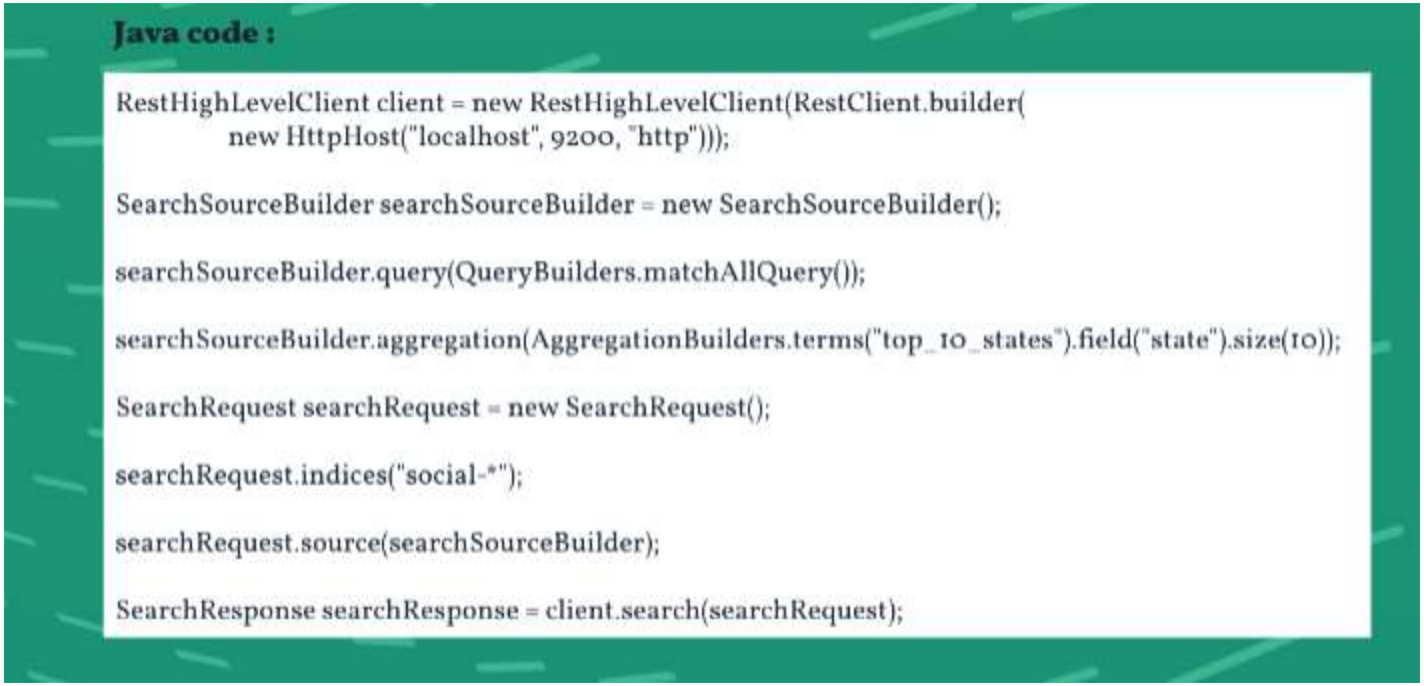
Java code :

```
Connection conn =
 DriverManager.getConnection("jdbc:elasticsearch:user=myuseraccount;password=mypassword;");

    Statement stat = conn.createStatement();

    ResultSet rs=stat.executeQuery("SELECT * FROM DocumentDB");

    while(rs.next()){
     for(int i=1;i<=rs.getMetaData().getColumnCount();i++)
     {
      System.out.println(rs.getMetaData().getColumnName(i) +"="+rs.getString(i));
     }
    }
```



Java code :

```
RestHighLevelClient client = new RestHighLevelClient(RestClient.builder(
        new HttpHost("localhost", 9200, "http")));

SearchSourceBuilder searchSourceBuilder = new SearchSourceBuilder();

searchSourceBuilder.query(QueryBuilders.matchAllQuery());

searchSourceBuilder.aggregation(AggregationBuilders.terms("top_10_states").field("state").size(10));

SearchRequest searchRequest = new SearchRequest();

searchRequest.indices("social-*");

searchRequest.source(searchSourceBuilder);

SearchResponse searchResponse = client.search(searchRequest);
```