What is Elasticsearch?

Elasticsearch is a distributed search and analytics engine built on Apache Lucene. Since its release in 2010, Elasticsearch has quickly become the most popular search engine and is commonly used for log analytics, full-text search, security intelligence, business analytics, and operational intelligence use cases.

On January 21, 2021, Elastic NV announced that they would change their software licensing strategy and not release new versions of Elasticsearch and Kibana under the permissive Apache License, Version 2.0 (ALv2) license. Instead, new versions of the software will be offered under the Elastic license, with source code available under the Elastic License or SSPL. These licenses are not open source and do not offer users the same freedoms. To ensure that the open-source community and our customers continue to have a secure, high-quality, fully open-source search and analytics suite, we introduced the OpenSearch project, a community-driven, ALv2 licensed fork of open-source Elasticsearch and Kibana.

How does Elasticsearch work?

You can send data in the form of JSON documents to Elasticsearch using the API or ingestion tools such as Logstash and Amazon Data 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. You can also use Kibana, a visualization tool, with Elasticsearch to visualize your data and build interactive dashboards.

Elasticsearch benefits

Fast time-to-value

Elasticsearch offers simple REST-based APIs, a simple HTTP interface, and uses schema-free JSON documents, making it easy to get started and quickly build applications for various use cases.

High performance

The distributed nature of Elasticsearch enables it to process large volumes of data in parallel, quickly finding the best matches for your queries.

Complimentary tooling and plugins

Elasticsearch comes integrated with Kibana, a popular visualization and reporting tool. It also offers integration with Beats and Logstash, helping you easily transform source data and load it into your Elasticsearch cluster. You can also use various open-source Elasticsearch plugins such as language analyzers and suggesters to add rich functionality to your applications.

Near real-time operations

Elasticsearch operations such as reading or writing data usually take less than a second to complete. This lets you use Elasticsearch for near real-time use cases such as application monitoring and anomaly detection.

Easy application development

Elasticsearch provides support for various languages including Java, Python, PHP, JavaScript, Node.js, Ruby, and many more.

Amazon Web Services (AWS) offers several services for deploying and managing Elasticsearch, a popular open-source search and analytics engine. Here are the key services and features related to Elasticsearch on AWS:

1. **Amazon OpenSearch Service (formerly Amazon Elasticsearch Service)**:
    - **Overview**: Amazon OpenSearch Service is a managed service that simplifies deploying, operating, and scaling OpenSearch and Elasticsearch clusters in the cloud. OpenSearch is a community-driven, open-source search and analytics suite derived from Elasticsearch 7.10 and Apache Lucene.
    - **Features**:
        - Fully managed clusters with automated backups, monitoring, and scaling.
        - Integrated with AWS Identity and Access Management (IAM) for secure access.
        - Supports both OpenSearch and legacy Elasticsearch APIs.
        - Advanced analytics features such as visualizations with OpenSearch Dashboards.
        - Support for log analytics, full-text search, and more.
2. **Amazon Managed Grafana**:
    - **Overview**: While not directly an Elasticsearch service, Amazon Managed Grafana is a fully managed service for Grafana, an open-source analytics and monitoring platform. It integrates seamlessly with Amazon OpenSearch Service for visualizing search and log data.
    - **Features**:
        - Managed Grafana dashboards for visualizing data from OpenSearch.
        - Secure and scalable with built-in integrations for various data sources.
3. **AWS Data Pipeline**:
    - **Overview**: AWS Data Pipeline is a web service that helps you process and move data between different AWS compute and storage services. It can be used to process logs or other data and feed it into an OpenSearch cluster for indexing and search.
    - **Features**:
        - Automated data workflows with built-in scheduling and error handling.
        - Integration with various AWS services, including Amazon S3 and Amazon RDS.
4. **AWS Lambda**:
    - **Overview**: AWS Lambda can be used in conjunction with OpenSearch to create event-driven processing pipelines. For example, you can use Lambda functions to process and transform data before indexing it into an OpenSearch cluster.
    - **Features**:
        - Serverless compute service that scales automatically.
        - Integration with other AWS services for event-driven architecture.
5. **Amazon Kinesis**:

- **Overview**: Amazon Kinesis provides real-time data streaming services. You can use it to stream data into OpenSearch for real-time analytics and search capabilities.
- **Features**:
  - Kinesis Data Streams for real-time data ingestion.
  - Kinesis Data Firehose for data delivery to OpenSearch.

6. **Amazon CloudWatch**:
   - **Overview**: Amazon CloudWatch is used for monitoring and managing AWS resources and applications. It can be used to monitor OpenSearch clusters and set up alarms based on metrics.
   - **Features**:
     - Provides metrics, logs, and alarms for OpenSearch clusters.
     - Integration with other AWS services for comprehensive monitoring.

These services can be used individually or together to build a scalable, secure, and highly available search and analytics solution on AWS.