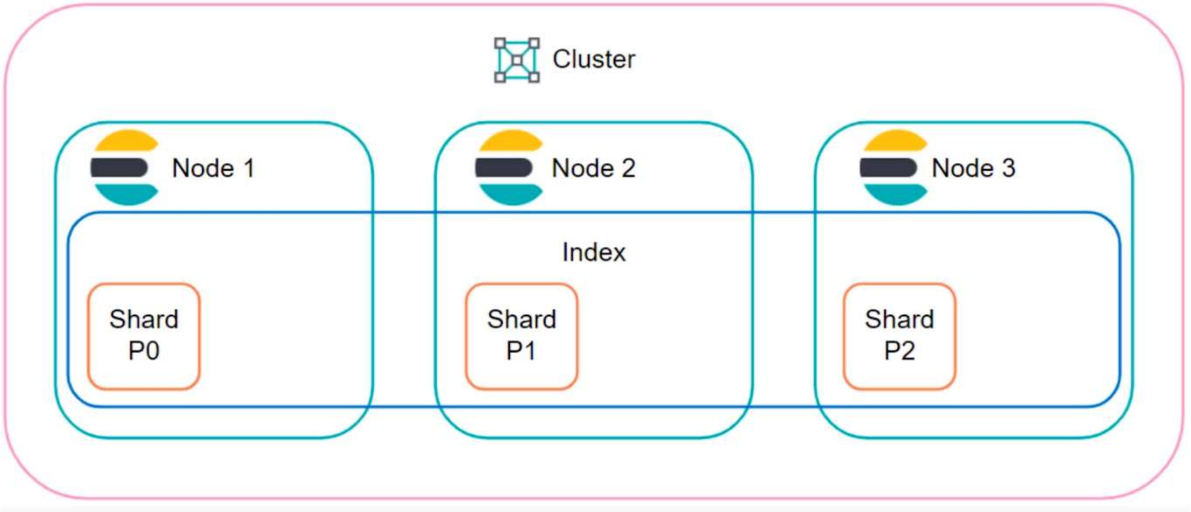
Elasticsearch Tech Review

Christian Gerber (cgerber3)

Elasticsearch (yes one word), is a search and analytics engine for searching all kinds of data. In short, it is a software that can help users find what they need faster, whatever the case. The types of data it can search are text, numerical, geospatial, structured, and unstructured data. The three main use cases it is used for is enterprise search, observability, and security. However, broken down, the use cases it can be used for are application searching, searching websites, intranet searches, log analytics, infrastructure metrics, performance monitoring, geospatial data analysis, security analysis, and business analytics. It is a near real-time system, the amount of time between a document being indexed until it is searchable is about one second. It is a powerful, developer friendly, tool, that can be used to improve the performance of your system.

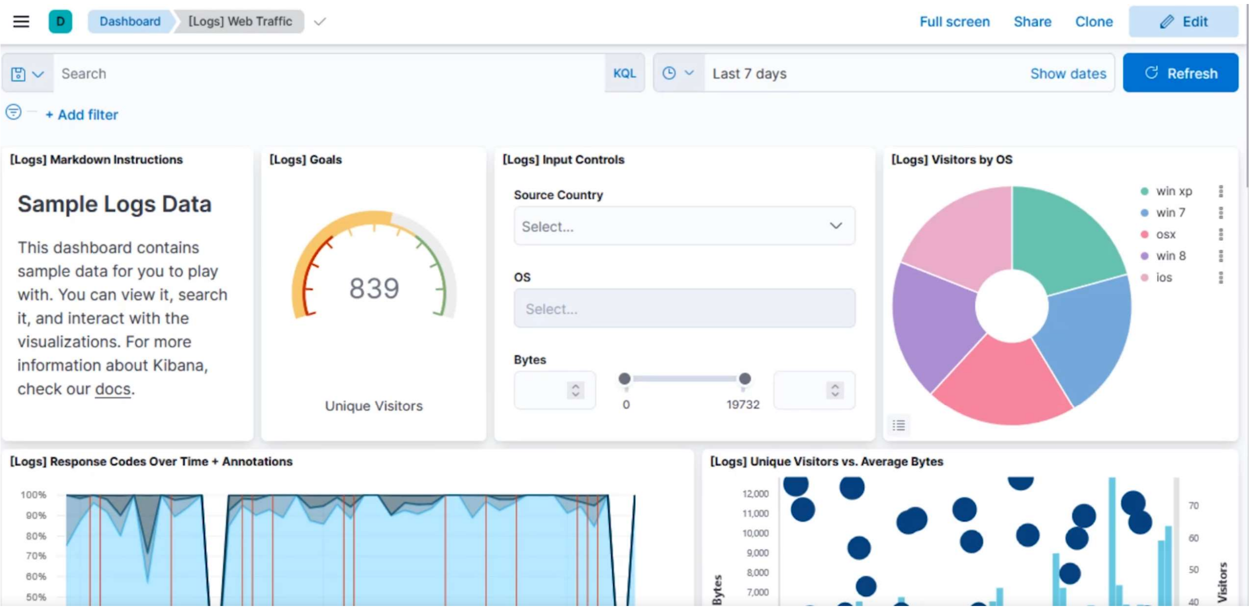
Elasticsearch works by ingesting data from many sources, it then parses the raw data, normalizes it, enriches it and then it is indexed in the system. The Elasticsearch index is a JSON document database of documents that are similar to each other. Each JSON document contains a set of keys, which are the names of the fields, and values, which is the data being stored. Elasticsearch does use an inverted index, which indexes every unique word and the number of times it appears in each document.

Elasticsearch is a distributed system, meaning that it has the ability to be scalable. A single instance of Elasticsearch is called a node, and a collection of nodes working together is called a cluster. This makes Elasticsearch fault tolerant because there is a dedicated “leader” node that tracks what all the other nodes are doing. Nodes are distributed into containers called shards, which are duplicated to provide redundancy of data. The shards are able to run in parallel, and often the search is distributed over these different shards, which enables Elasticsearch to process a large volume of data and quickly find the match it is looking for. These shards can be housed in hundreds of servers which allows for ingestion of petabytes of data.



Example of Elasticsearch architecture.

Elasticsearch tightly integrates two other technologies into it to become the Elastic Stack. These two other technologies are Logstash and Kibana. Logstash is used to send data to Elasticsearch. It will ingest data from multiple data sources at the same time and transform it to be put into Elasticsearch. Kibana is a tool that visualizes data for Elasticsearch, it is the “window” that you can peer through into the system. It provides real time graphs that can be used for analyzation as well as actionable items such as webhooks or the ability to send emails for when you are using Elasticsearch to monitor a system. You can also use it as an interface, to specifically search for items in the dashboard



Sample Logs Data

This dashboard contains sample data for you to play with. You can view it, search it, and interact with the visualizations. For more information about Kibana, check our docs.

Example of Kibana Dashboard

Elasticsearch is also a great tool because it supports many different types of deployments and many different types of programming languages. You can either deploy it on your own hardware, or Amazon Web Services, Google Cloud and Microsoft Azure all have a service that can be installed to run it on there. There is also a subscription you can pay for Elastic Cloud Enterprise which allows you to manage everything from a single point, but not deploy it on a public cloud. Elasticsearch also supports interfacing through Java, JS, Go, .NET, PHP, Perl, Python, Ruby, and REST API.

Overall, Elasticsearch is a great tool to increase the performance of searching for text information in your software. With the scalability, versatility, and performance of the software, there is almost no reason not to use it in your application. Searches will become more accurate, and faster within your system.

Works Cited

Amazon. “What Is Elasticsearch.” *Amazon*, Amazon, 2022, https://aws.amazon.com/what-is/elasticsearch/.

Elastic. “What Is Elasticsearch?” *Elastic*, Elastic, 2022, https://www.elastic.co/what-is/elasticsearch.

Kobar, George. “Getting Started with Elasticsearch.” *Elastic*, 20 June 2022, https://www.elastic.co/webinars/getting-started-elasticsearch?elektra=en-whatis-elasticsearch.