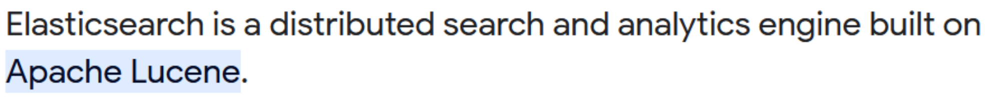
1. What is Elasticsearch?

Answer: Elastic search is Apache lucent based search engine for logs, and NoSQL DB.



Elasticsearch is a distributed search and analytics engine built on Apache Lucene.

Lucene Core is a Java library providing powerful indexing and search features, as well as spellchecking, hit highlighting and advanced analysis/tokenization capabilities. The PyLucene sub project provides Python bindings for Lucene Core.

2. What is Elasticsearch used for?

Answer: For observability.

3. What is an Elasticsearch cluster?

Answer: An Elasticsearch cluster is a group of node instances that are connected together, a cluster can also consist of a single node.

4. What is an Elasticsearch Index?

Answer: An Elasticsearch index functions as a database containing different defined types of data for easier organisation and retrieval, its indexes map to a primary shard/s.

5. What does the term document mean in Elasticsearch?

Answer: Documents are JSON objects that are distributed across your Elasticsearch cluster which are accessible from any node.

6. What is ELK?

Answer: ELK (also known as the Elastic Stack) stands for the formally open-source tools Elasticsearch, Logstash, and Kibana.

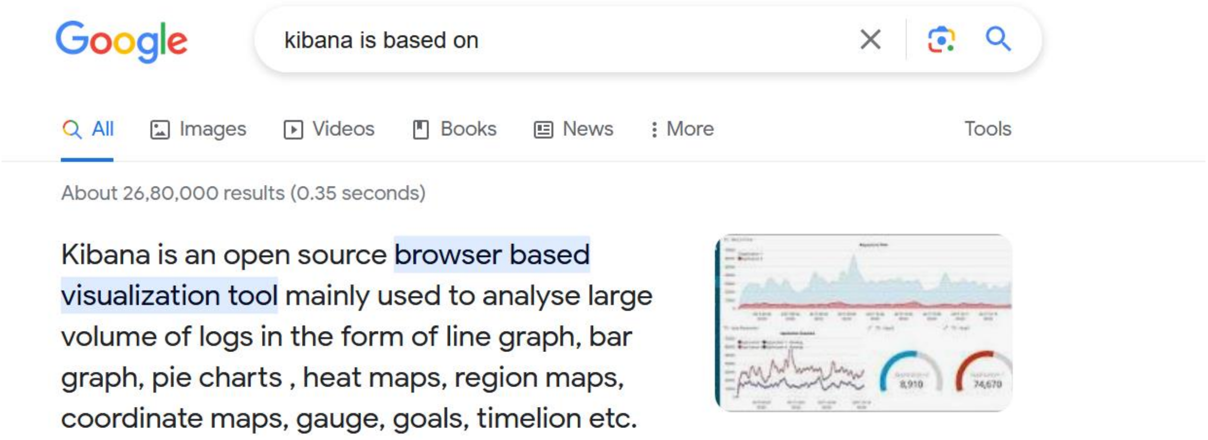
7. What is Logstash?

Answer: Logstash is a powerful, flexible pipeline that collects, enriches and transports data. It works as an extract, transform & load (ETL) tool for collecting log messages and forwarding them onto Elasticsearch for visualisation within Kibana.

8. What is Kibana?

Answer: Kibana is an open-source visualisation and reporting user interface developed in 2013 by Rashid Khan of the Elastic Company.

Kibana allows users to produce visualisations from a variety of data inputs & can create pie charts, heat maps, line graphs, scatter plots and so much more to reflect data in an easily digestible format.



About 26,80,000 results (0.35 seconds)

Kibana is an open source browser based visualization tool mainly used to analyse large volume of logs in the form of line graph, bar graph, pie charts , heat maps, region maps, coordinate maps, gauge, goals, timelion etc.

9. What are the advantages of using Elasticsearch?

Answer: Some of the key benefits users of Elasticsearch commonly cite include:

- Lightning-fast performance even when working with massive-scale datasets
- Its ability to scale
- Extensive API
- Multilingual

10. What are the advantages of using Logstash?

Answer: The key features that users of Logstash find beneficial include:

- Over 200 plugins available
- Process unstructured data
- Pre-built and custom filters
- Built custom data processing pipelines
- Works as an extract, transform & load (ETL) tool

11. What are the advantages of using Kibana?

Answer: A few of the main benefits of using Kibana are as follows:

- Real-time observability
- Integration with Elasticsearch
- Browser-based visualisation tool
- Many graphs and charts to select from

Human mind works in image and it’s good to visualize data.

12. What are the advantages of using the ELK Stack?

Answer: The ELK Stack can be used for a wide variety of use cases including but not limited to;

- APM
- SIEM
- Log analysis

- Server monitoring
- Container monitoring
- Metrics management
- Vulnerability scanning
- Compliance & auditing
- Infrastructure monitoring
- Monitoring website uptime
- Measuring sales performance
- Understanding user behaviour

13. How do you check the version of Elasticsearch you are working with?

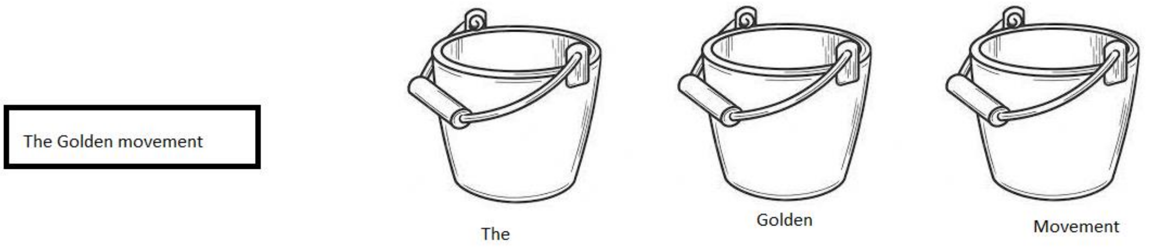
Answer: To find out which version of Elasticsearch you are running locally execute the following curl command from your command line: curl -XGET 'http://localhost:9200'

14. Is Elasticsearch & Kibana considered open-source?

Answer: The last release of Elasticsearch and Kibana classed as open-source is version 7.10 as the newly launched 7.11 does not meet the Open-Source Initiative's requirements for open-source licensing due to the terms detailed within Elastic's Server Side Public License (SSPL).

15. What is bucketing in Elasticsearch and Kibana?

Answer: Bucket aggregations create buckets of documents. In our example of a red car, an aggregation on the field will return a "red" bucket and a "car" bucket. Any document with a mention of the word red in this text field will be added to the "red" bucket and the same for the word car and the "car" bucket. Obviously, if we have lots of words in a text field over many documents we will end up with a lot of buckets. Some documents will be found in more than one bucket depending on the content of the field whilst others may not.



16. How do you create an index in Elasticsearch?

Answer: Use the following command to create a new Elasticsearch index: PUT /my-index-000001

17. How do you load data into Elasticsearch?

Answer: To get started with loading data into Elasticsearch you will ideally want to use one of the many available Beats (Elastic's own log shippers) to load data into Logstash for further processing within Elasticsearch. Popular Beats include the following:

- Filebeat: Lightweight shipper for logs
- Metricbeat Collect metrics from your systems and services.
- Heartbeat specify the protocols to monito
- Auditbeat *Auditbeat* allows you to monitor user activity and processes and analyze your event data in the Elastic Stack without touching auditd
- Packetbeat *Packetbeat* is a lightweight network packet analyzer that sends data from your hosts and containers to Logstash or *Elasticsearch*.
- WinLogBeat he open source tool for shipping Windows event logs to Elasticsearch to get insight into your system

18. Where does Elasticsearch store data?

Answer: You can locate where Elasticsearch stores data under its default paths. For Debian/Unbuntu this will be located at /var/lib/elasticsearch/data & for RHEL/CentOS this will be located at /var/lib/elasticsearch.

19. How do you check if Elasticsearch is running?

Answer: To verify Elasticsearch is running on Linux you will want to enter the following curl command using the command line: curl -XGET http://localhost:9200/_status

20. How do you stop Elasticsearch?

Answer: To stop the Elasticsearch service on Linux you will want to change the directory in a terminal window to ES_HOME/bin. and use kill to stop Elasticsearch search, you can find the process ID (pid) by using the following command ps -ef | grep elas.

21. What is AWS Elasticsearch?

Answer: AWS Elasticsearch (also known as Open Distro for Elasticsearch) is currently being rebranded to Open Search & Open Dashboards. AWS's offering of Elasticsearch & Kibana has a number of additional features including trace analytics, scheduled reports, index, document and field level security as well as a SQL and PPL query workbench.

22. What database does Elasticsearch use?

Answer: Elasticsearch is a NoSQL database

23. How many indexes does Elasticsearch have?

Answer: Elasticsearch does not impose strict limits on the number of indexes you can have but having too many indexes can affect how easily you can scale. As each index could easily handle in excess of 20GB you should consider why you might require so many separate indexes.

24. How can you determine how much storage is being used by Elasticsearch?

Answer: In Linux, you can view how much space your locally hosted Elasticsearch setup is using by entering the following command from the command line: du -hs /myelasticsearch/data/folder

25. What is the curl command?

Answer: The curl command in Elasticsearch allows you to take many actions including but not limited to: deleting indexes, list all indexes, list all documents within an index, query using URL parameters, add data and list index mappings.

26. How do you delete an index in Elasticsearch?

Answer: By using the DELETE /index name. Command.

27. How do you install Logstash plugins?

Answer: To find individual download and install commands specific to the Logstash plugin you are wishing to install, RubyGems provides a wealth of resources to help you. Once you've selected the plugin you wish to use you can add it to your Logstash installation using the following command: bin/logstash-plugin install logstash-input-github

28. Where can you write Logstash configuration in Elasticsearch?

Answer: Logstash settings can be found located at the {extract.path}/config directory. Logstash settings can be configured from the following files found here: logstash.yml, pipelines.yml, jvm.options, log4j2.properties and within Linux there is an additional configuration file called startup.options.

29. How can you use Logstash GeoIP?

Answer: The Logstash GeoIP filter is used to add supplementary information on the geographic location of IP addresses based on data available within the MaxMind GeoLite2 database.

30. How can you forward logs from Kubernetes to Logstash?

Answer: By using a log shipper such as Filebeat, as illustrated on our integration page for sending Kubernetes logs to Logstash.

For additional resources on tools suited to K8s then check out our guide on Kubernetes management tools.

31. How can you test Logstash performance?

Answer: You can use the node stats API to retrieve runtime statistics from Logstash.

32. What is a Logstash pipeline?

Answer: A Logstash pipeline consists of these elements as listed in the following order: Input (the source of data), filter (processing, parsing & enrichment of data) & output (write the data for use within Elasticsearch & Kibana).

33. How do you start Kibana?

Answer: When Kibana has been installed using a .tar.gz package on Linux, it can be started from the command line using the following command: ./bin/kibana For additional installation types & operating systems consult the following guide.

34. How do you search Kibana?

Answer: In Kibana you can easily search the current index pattern you are viewing by entering your search criteria in the Query bar visible near the top left-hand side of your screen. Conveniently Kibana also shows its users a filtering dialogue which provides somewhat of a cheat sheet for easier filtering to be conducted.

35. How can you create a Kibana dashboard?

Answer:Once you have Kibana loaded you will want to open the main menu and select Dashboard, from here you can select Create Dashboard. Once this step has been completed you will need to add panels of data to your dashboard for further visualisations and chart type selection to be applied.

36. Where are Kibana dashboards stored?

Answer: Kibana dashboards are stored in Elasticsearch under the default index kibana-int which can be edited within the config.js file if you wish to store your dashboards in a different location.

37. What are common reasons that Kibana might be slow to load?

Answer: If you find that your Kibana instance is loading slowly it is often mentioned in the support forums that the reason this happens is due to the Apps bundle or apps themselves loading in the background.

38. What is the line chart used for in Kibana?

Answer: A line chart (also known as a line graph) is a type of data visualisation that displays data as a series of points that reflects changes over a designated time period.

39. What is Filebeat?

Answer: Filebeat is the leading choice for forwarding logs to the Elastic Stack due to its reliability & minimal memory footprint. Filebeat was originally written in the Go programming language and its features originated from a combination of the best attributes of Logstash-Forwarder & Lumberjack. Additionally, when Filebeat is part of the logging pipeline it can generate and parse common logs to be indexed within Elasticsearch. You may often see Filebeat mentioned alongside Logstash as the two are used in tandem with each other for the majority of logging use cases.

40. What is Metricbeat?

Answer: Metricbeat is a metrics shipper built on the Libbeat framework. It originated from Topbeat (which has now been deprecated) and is primarily used for collecting metrics prior to their enrichment within Logstash for further processing within Elasticsearch & Kibana. Some users of Metricbeat may not wish to automatically push their metrics data to Logstash, in this instance they would likely use a service (for example Kafka or Redis) to buffer the data.

41. What is Journalbeat?

Answer: Journalbeat is one of the most recent additions to the Beats family. This particular Beat is used to collect log entries from Systemd Journals. As with the other Beats, Journalbeat is based on the libbeat framework. Journalbeat is rated as being easier to use than more commonly known Beats for collecting Systemd Journal logs (such as Filebeat) but is regarded as an experimental Beat so may be subject to change.

42. What is Heartbeat?

Answer: Heartbeat is a lightweight shipping agent that was created to allow observability of the health of services running on a specified host, its results can then be forwarded to Logstash for further processing. Heartbeat is notable for the fact that it is the only member of the Beats family that Elastic themselves recommend you to install on a separate network/machine external to the one you are currently wishing to monitor.

43. What is Packetbeat?

Answer: Packetbeat is a network package analyser used to capture network traffic. Packetbeat can be used to extract useful fields of information from network transactions before shipping them to one or more destinations, including Logstash. This is especially useful for those that wish to troubleshoot and detect performance hits.

44. What is WinLogBeat?

Answer: Winlogbeat is a log shipper used for collecting Windows event logs as it can easily read events from any event log channel using the Windows operating system. Windows log data once centralised within the ELK stack can then be monitored for anomaly detection & other security-related incidents.

45. What is Auditbeat?

Answer: Auditbeat is a lightweight shipper used to collect audit data from your systems. This Beat can also be used to detect crucial and unexpected changes to configuration files & binaries which can often be vital for pinpointing compliance issues and security violations occurring within your organisation.

46. What is reindexing?

Answer: Reindexing your Elasticsearch index is mainly required in the event that you wish to update mapping or settings associated with your current index. Reindexing means that you are copying preexisting data from an index that already exists to a new destination index. The command endpoint _reindex can be used for this purpose.

47. How can you make balanced shards?

A: Your clusters and/or shards are considered balanced when they have an equal number of shards across each node, thankfully Elasticsearch will run an automatic process of rebalancing shards which moves shards between the nodes that make up your cluster in order to improve its performance. You may need to take manual action if your configurations for forced awareness or allocation filtering clashes with Elasticsearch's attempts to automatically rebalance shards.

48. What is Grok?

Answer: Grok is a filter plugin for Logstash that is used to parse unstructured data. It is often used for transforming Apache, Syslog and other webserver logs into a structured and queryable format for easier data analysis to be performed.

49. What is fuzzy search?

Answer: Fuzzy search allows Elasticsearch to return positive matches for non-exact matches of the term you are searching for. This is especially beneficial for eCommerce retailers where site visitors may often have typos in their spelling when trying to locate a product they wish to purchase. Fuzzy match results mean that these visitors are not served with a blank page which would often lead to a user being less likely to convert.

50. How do you view an index template?

A: There are two ways you can view the current index template of your ELK Stacks, one of these uses Kibana dev tools and the second option involves using the Template Index API.

50. What is Apache lucent and how it's related to ELK?

Answer: There are two ways you can view the current index template of your ELK Stacks, one of these uses Kibana dev tools and the second option involves using the Template Index API.