**Day2: Understanding of ELK stack**

This 30 day challenge will make use of the ELK stack to create a fully functioning SOC environment. To accomplish this, a good understanding of the components that will make this possible is crucial.

**Telemetry : An Overview**

Telemetry refers to the automated process of collecting, transmitting, and analyzing data from remote sources or devices to monitor their performance, behavior, or condition. In cybersecurity and IT, telemetry is used to gather information about systems, networks, and applications to detect anomalies, ensure security, and optimize performance.

**How is Telemetry Collected?**

Two of the most popular ways to collect telemetry are:

**Beats :** Beats are lightweight data shippers that collect and send data from various sources to Logstash or Elasticsearch. Depending on the type of telemetry, the required beat is installed on the endpoint. Each Beat is specialized for a specific type of data, such as:

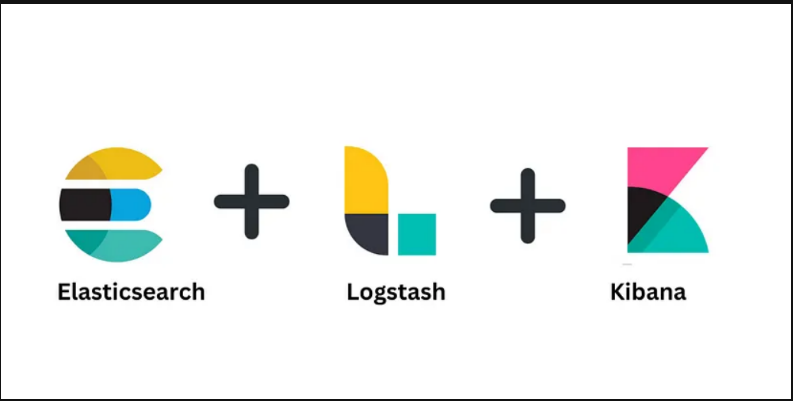
* Filebeat - for log files
* Metricbeat - for system metrics
* Packetbeat - for network data
* WinLogbeat - for Windows event logs
* Auditbeat - for audit data
* Heartbeat - for uptime monitoring

**Elastic Agents**

Elastic Agents are unified, lightweight agents that collect and ship data to the Elastic Stack. They are designed to replace multiple Beats with a single, versatile agent that can be centrally managed using Fleet in Kibana. Elastic Agents simplify data collection across systems, networks, and endpoints by handling logs, metrics, security data, and more, providing a streamlined approach to monitoring and security operations. For this challenge, I will be using an Elastic Agent.

**What is the ELK Stack?**

The ELK Stack consists of **Elasticsearch**, **Logstash**, and **Kibana**. It is a powerful open-source toolset for searching, analyzing, and visualizing log data in real-time. Let's analyze these 3 components in detail:



**E : Elasticsearch**

Elasticsearch is a distributed, open-source search and analytics engine designed for handling large volumes of data in real-time. It is the core component of the Elastic Stack (ELK Stack), used for indexing, searching, and analyzing data. It is widely used for use cases such as log and event data analysis, full-text search, monitoring, and business intelligence. Elasticsearch uses the ES | QL (Elasticsearch Query Language to perform queries on databases.

**Key Features of Elasticsearch:**

1. **Full-Text Search**: Allows efficient and scalable text search, making it ideal for applications requiring quick search capabilities.
2. **Distributed Architecture**: Scales horizontally by distributing data across multiple nodes, providing high availability and fault tolerance.
3. **Real-Time Analytics**: Supports near real-time search and data analysis, enabling users to gain insights from large datasets almost instantly.
4. **Scalability and Flexibility**: Easily scales up or down depending on data volume and query load, with support for a wide variety of data types and structures.
5. **RESTful API and JSON**: Uses a RESTful interface for data manipulation and search operations, making it accessible and easy to integrate with other applications and systems. Using JSON and Restful APIs also makes sure that many different applications can interact with our databse in a programmable way.

**L : Logstash**

Logstash is an open-source data processing pipeline that is part of the Elastic Stack (ELK Stack). It collects, processes, and transports data from various sources to destinations like Elasticsearch, making it an essential tool for data ingestion and preparation. It is commonly used for centralizing and processing logs, metrics, and events in IT environments, providing the backbone for robust logging and monitoring solutions.

**Key Features of Logstash:**

1. **Data Ingestion**: Supports a wide range of input sources, including logs, metrics, and other data types from various platforms, such as servers, databases, and applications.
2. **Data Processing**: Provides a rich set of filters and plugins for parsing, transforming, and enriching data. This includes operations like field extraction, data normalization, geo-location, and more.
3. **Extensibility**: Highly extensible through numerous plugins, allowing users to customize and extend its functionality for different data processing needs.
4. **Flexible Output**: Can output data to multiple destinations, including Elasticsearch, databases, file systems, and other data stores, making it versatile for various use cases.
5. **Real-Time Processing**: Processes and transforms data in real-time, ensuring timely data availability for analysis and monitoring.

**K : Kibana**

Kibana is a data visualization and exploration tool that is part of the Elastic Stack. It provides a web-based interface for visualizing data stored in Elasticsearch, allowing users to create dynamic dashboards, graphs, and reports. Kibana is widely used for analyzing log and time-series data, monitoring performance, and supporting cybersecurity operations by enabling quick insights into large datasets through its powerful querying and visualization capabilities.

**Key Features of Kibana:**

1. **Data Visualization and Dashboards**: Create a variety of visualizations and combine them into interactive dashboards for comprehensive data analysis and monitoring.
2. **Search and Query**: Use advanced query languages to search, filter, and explore data efficiently, making it easy to find specific insights.
3. **Machine Learning and Alerts**: Detect anomalies, predict trends, and set up alerts to monitor data and receive notifications based on defined conditions.
4. **Maps and Geospatial Analysis**: Visualize and analyze geographic data using interactive maps, supporting multiple layers and spatial exploration.
5. **Security and Access Control**: Implement role-based access controls to manage user permissions and protect sensitive data within the platform.



**Benefits of the ELK Stack**

1. **Centralized logging:** It allows one to meet compliance requirements and search through data in case of a security incident.
2. **Flexibility:** Using Elastic Agents or Beats, it allows one to ingest data and logs into Elastic instance or Logstash as per our own requirements.
3. **Visualization:** Using customizable dashboards, one can have all important data available at a glance.
4. **Scalability:** The ELK stack can be configured to accommodate and handle larger environments.
5. **Ecosystem:** The ELK stack is compatible with a wide variety of cybersecurity tools, and has an active and growing community that one can reach out to, for support.

**Conclusion**

With invaluable guidance from Mr. Stevens at MYDFIR ([his website](https://www.mydfir.com/)) and his [YT video](https://www.youtube.com/watch?v=4AwBhXAW90Q&list=PLG6KGSNK4PuBb0OjyDIdACZnb8AoNBeq6&index=3) outlining day 2 of the 30-Day SOC Challenge, I learned about the ELK stack, its components, functionalities, features and advantages. I look forward to using the knowledge I gained today as part of this SOC project as well as beyond it in my cybersecurity career.