

This system architecture diagram illustrates the complete pipeline for **Advanced Persistent Threat (APT) Detection** using the **ELK Stack** (Elasticsearch, Logstash, and Kibana) within your organization (NTRO).

**1. Data Sources**

This section represents the origin of the raw data that needs to be analyzed for signs of an APT.

* **Endpoint Logs (Servers & Cloud Logs):** Logs generated by operating systems, applications, and security tools on individual computers and servers (e.g., Windows Event Logs, Sysmon data, Linux Audit logs). These logs are crucial for detecting actions like **Credential Dumping** and **PowerShell Execution**.
* **Network Traffic Logs:** Data from network devices (firewalls, routers, proxies) and network monitoring tools. This is key for identifying **Lateral Movement** and **Unusual Egress Data**.
* **APT Sample (Log Generation):** A crucial part of your project. A known APT malware sample is executed in a controlled, isolated environment to generate **realistic, malicious log data**. This data is then used to test and validate the detection rules.

**2. Ingestion & Processing (Logstash)**

**Logstash** is the data processing pipeline of the ELK Stack.

* **Function:** It takes the raw logs from the data sources, processes them, and prepares them for storage.
* **Processes:** **Data Processing, Filtering, and Enrichment** (e.g., parsing raw text logs into structured fields like IP addresses, usernames, and event types; normalizing data to a common schema like the **Elastic Common Schema (ECS)**).

**3. Data Storage & Indexing (Elasticsearch)**

**Elasticsearch** is the central, distributed search and analytics engine.

* **Function:** It efficiently stores, indexes, and allows for near real-time searching of the vast amounts of processed log data.
* **Key Data Points:** The logs are indexed here, allowing analysts to search for specific APT activities like **Credential Dumping, Internal Reconnaissance, PowerShell Execution**, and **Unusual Egress Data**.

**4. Visualization & Analysis (Kibana)**

**Kibana** is the user interface for the ELK Stack.

* **Function:** It allows users to visualize the data, run queries, and, most importantly, **develop and manage Threat Rules**.
* **Threat Rules Development (APTs):** This is the core focus of your project. Rules are developed within Kibana (often using the built-in **Elastic Security** features like Detection Engine) to look for specific sequences, volume changes, or patterns in the indexed data that match known **APT TTPs** (Tactics, Techniques, and Procedures).

**5. Security Operations Center (SOC)**

This represents the final output and operational stage of the detection system.

* **APT Attack Lifecycle Detection:** The developed threat rules successfully detect the various **stages of an APT attack**, such as:
  + **Reconnaissance** (e.g., unusual repeated failed logins).
  + **Initial Access** (e.g., a specific payload execution).
  + **Persistence** (e.g., new startup entry creation).
  + **Credential Dumping**.
  + **Lateral Movement** (e.g., remote logins from unusual systems).
  + **Data Exfiltration** (e.g., large transfer of data to an external, known-bad IP).
* **Security Analysts:** Personnel monitor the alerts generated by Kibana.
* **Alerting & Incident Response:** Upon rule activation, an alert is triggered, leading to necessary **incident response** actions to contain and eradicate the detected APT.