Logging with ELK Stack (Elasticsearch, Logstash, Kibana)

Introduction

Modern applications generate massive amounts of log data. To effectively collect, parse, store, and visualize this data, the **ELK stack**—comprising **Elasticsearch**, **Logstash**, and **Kibana**—is one of the most popular open-source solutions used in production environments.

Overview of ELK Stack

1. **Elasticsearch** A distributed, RESTful search and analytics engine capable of storing large volumes of data. It stores the logs and provides powerful full-text search capabilities.

2. **Logstash** A data processing pipeline that ingests logs from multiple sources, transforms them, and sends them to a destination (usually Elasticsearch). It supports a wide range of input, filter, and output plugins.

3. **Kibana** A web-based UI for visualizing data stored in Elasticsearch. It allows users to create dashboards, search logs, and monitor application behavior.

Setup and Workflow

1. Log Collection with Filebeat

Install Filebeat on your servers to monitor log files (e.g., `/var/log/syslog`, `/var/log/nginx/access.log`). Filebeat forwards log entries to Logstash or Elasticsearch.

```
# filebeat.yml
filebeat.inputs:
  - type: log
    paths:
      - /var/log/*.log
output.logstash:
  hosts: ["localhost:5044"]
```

2. Log Processing with Logstash

Logstash processes incoming logs with a configurable pipeline of inputs, filters, and outputs.

```
# Example Logstash config: logstash.conf
input {
  beats {
    port => 5044
  }
}
filter {
  grok {
    match => { "message" => "%{COMMONAPACHELOG}" }
  }
  date {
    match => ["timestamp", "dd/MMM/yyyy:HH:mm:ss Z"]
  }
}
output {
  elasticsearch {
    hosts => ["http://localhost:9200"]
    index => "weblogs-%{+YYYY.MM.dd}"
  }
}
```

3. Log Storage with Elasticsearch

Once Logstash sends structured logs to Elasticsearch, they are indexed and stored. Elasticsearch allows fast querying using a powerful DSL (Domain-Specific Language).

4. Log Visualization with Kibana

Connect Kibana to Elasticsearch, and you can:

- Search logs by keyword, fields, timestamps
- Create dashboards for real-time monitoring
- Set up alerts and anomaly detection

Key Benefits

- Scalable and distributed architecture
- Rich querying and full-text search
- Real-time visualization
- Ecosystem integrations (e.g., Beats, APM agents)

Conclusion

The ELK stack is a powerful solution for managing and analyzing logs from various sources in real time. With its modular components and strong community support, it's widely adopted for observability, troubleshooting, and monitoring needs across industries.