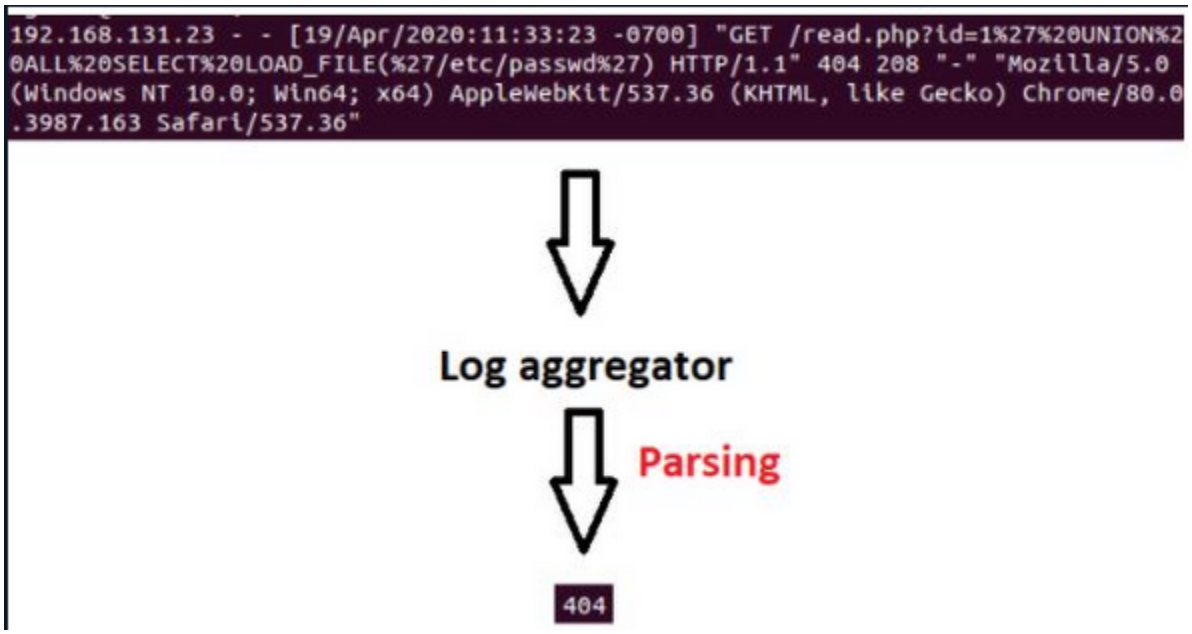
Log Aggregation and Parsing Analysis

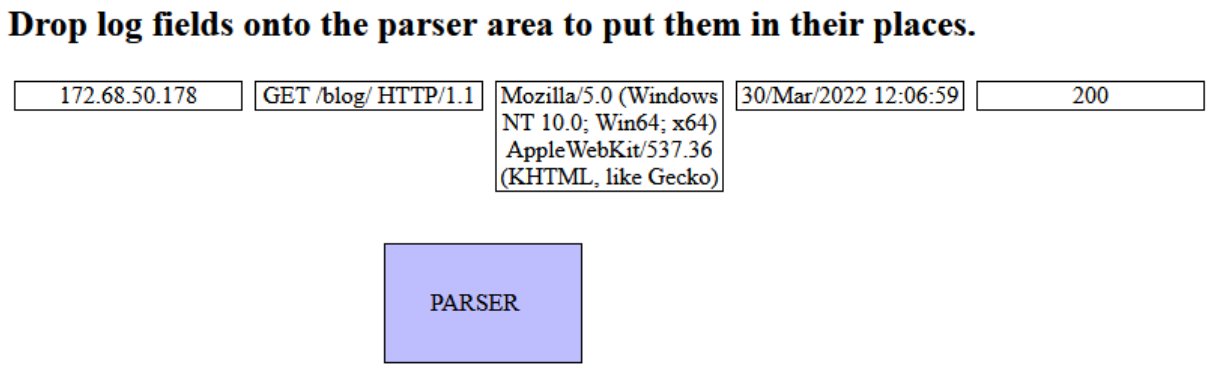
By: Ryan Stewart

Log Aggregation and Parsing

The initial destination for the generated logs is the log aggregator, where logs can undergo editing before being forwarded to their final destination. For instance, if specific information, such as status codes from web server logs, is required, the incoming logs can be filtered, sending only the relevant portions to the designated target.



Aggregator EPS

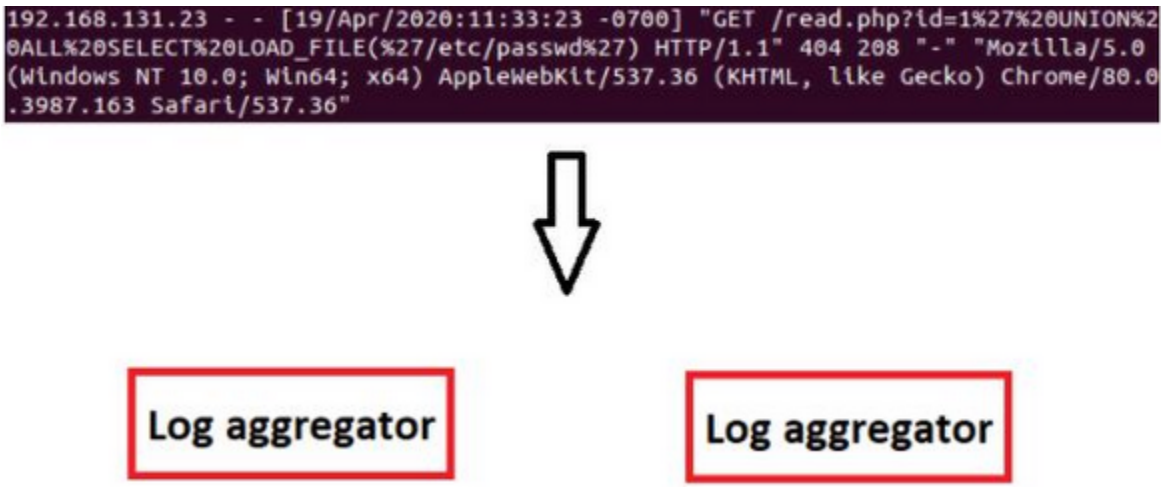


What is EPS?

EPS stands for Events Per Second, calculated as Events divided by the time period in seconds. For instance, if the system receives 1000 logs in 5 seconds, the EPS would be 1000/5 = 200. As the EPS value increases, the resources required for the aggregator and storage also increase.

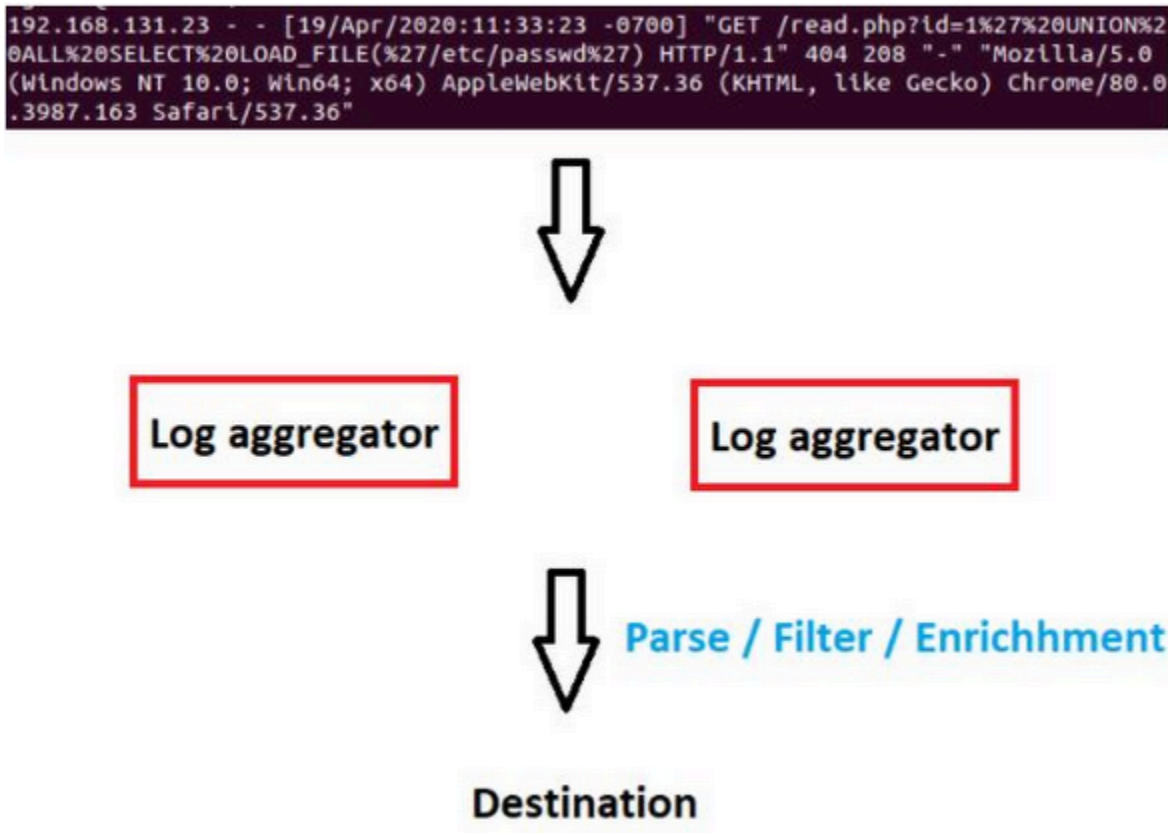
Scaling the Aggregator

To distribute the load evenly and prevent overloading a single aggregator, multiple aggregators can be implemented. These aggregators can be selected in either a sequential or random manner.



Log Aggregator Process

Incoming logs to the aggregator undergo processing before being directed to the target. This process may involve parsing, filtering, and enrichment.



Log Modification

In certain scenarios, adjustments may be necessary for incoming logs. For example, if date information in most logs follows the format dd-mm-yyyy, but a particular source provides mm-dd-yyyy, **conversion of the date format is essential.** Another instance could involve **converting** incoming time information from UTC + 2 to UTC + 1.

Log Enrichment

Enrichment is performed to **enhance the effectiveness** of collected logs and save time. Examples of enrichments include:

1. Geolocation: Determining and adding the geolocation of a specified IP address to the log, facilitating location-based analysis and saving time for log viewers.

2. DNS: Conducting DNS queries to find the IP address of a domain or performing reverse DNS to find the domain associated with an IP address.

1. What is the Events Per Second (EPS) of a SIEM system that receives 150,000 logs per minute? 150000/60 =2500

2. Among the following options, which is not a skill of a log aggregator?

- Filtering
- Parsing
- Analysis
- Enrichment