Augtera The First Network AI for the Modern Enterprise

Solution Brief

LogAI Integration with Splunk and Elastic Search

Introduction

Logs are a rich source of network telemetry information for network operations. While there are several tools that exist for historical analysis and querying of logs, they have not provided real-time anomaly detection capabilities that leverage the richness of text-based logs. This is increasingly relevant as operators want to find needles automatically and proactively in the growing haystack of log data. In addition, the volume, velocity, and variety of log messages calls for new approaches that generate structure and efficient analysis from unstructured text messages.

LogAI was developed to address Network Operations uses cases. To process in real-time, streaming log data, from any log, without latency or message drops. Specifically, to address the challenges of producing actionable insights from unstructured text data at such high volumes, where noise levels are so high, and the unknown unknowns are prevalent and undetected. This solution brief will describe how LogAI integrates with Splunk and ElasticSearch to enable seamless deployment in existing Enterprise ecosystems.

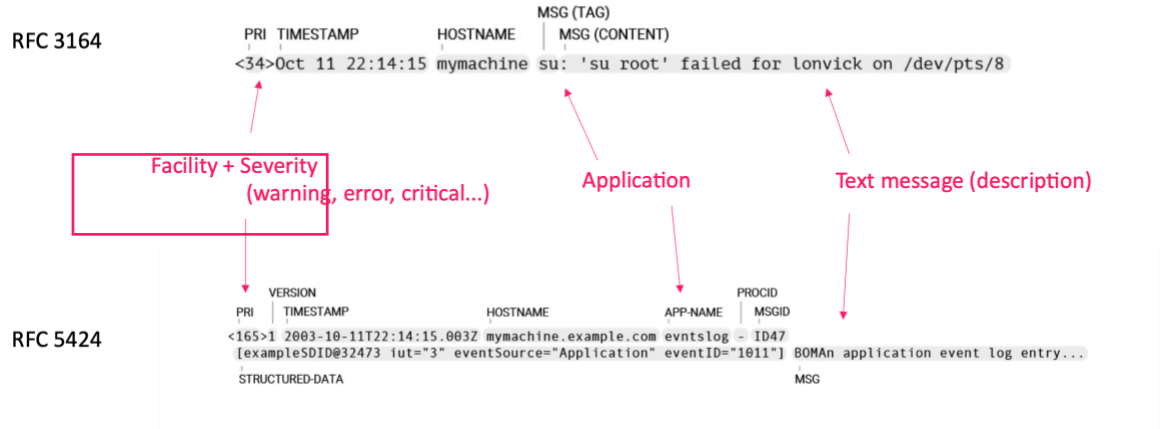
LogAI Flexible Data Ingestion

Syslog has long been the standard of network equipment vendors, however with the rise of cloud-based systems and message streaming technologies new logs formats, such as encoding logs in JSON are becoming more common. However, logs are generated by Cloud based systems as well as equipment vendors in many other non-standard formats, typically encoded as JSON. Further logs are collected, normalized, and distributed by operations teams in many other

formats, typically also encoded as JSON. Kafka is an emerging message bus for many types of data within Network Operations environments.

LogAI supports all these scenarios today and at its core is agnostic to the format of the log messages when they are ingested, as all logs are normalized to a common internal format.

Image 1: Structure of a Syslog message, one of the log format supported by Augtera



Logs can be ingested directly from devices to Augtera platforms, or forwarded by existing customer log tools such as Splunk or ElasticSearch.

Log Ingestion from Splunk Forwarders

Splunk forwarders can forward raw data to non-Splunk systems packaged in standard syslog.

The syslog output processor sends RFC 3164-compliant events to a TCP/UDP-based server and port, making the payload of any non-compliant data RFC 3164-compliant.

By default, Splunk software does not change the content of an event to make its character set compliant with Augtera Network AI platform.

The main configuration piece is done in file `$SPLUNK_HOME/etc/system/local/outputs.conf` where Augtera Network AI platform IP address is configured as well as udp port 514 as a destination.

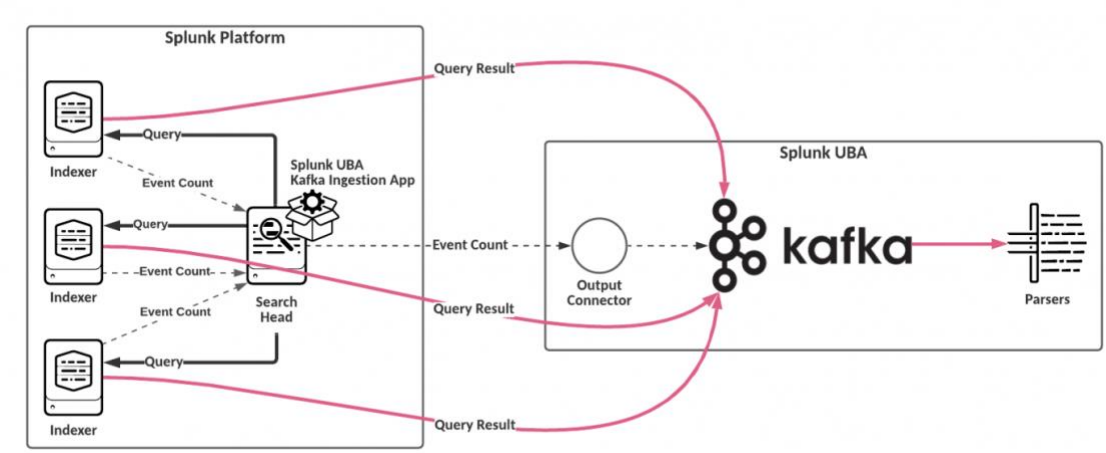
It is also possible to configure Splunk to send a subset of data to Augtera, for example only the data from hosts whose names begin with "nyc". You have to edit files `props.conf` and `transforms.conf` to specify the filtering criteria.

Send Data from the Splunk Platform directly to Kafka

When working with large data sets, you can send events from the Splunk platform directly to Kafka for ingestion. Sending data directly to Kafka offloads the processing task from the search heads to the indexers. The search heads still track the total number of events processed.

After Kafka ingestion is enabled, events from your search results are pushed from the indexers directly to Kafka on Splunk UBA.

Image 2: push data from the indexers directly to Kafka in Splunk UBA



Log Ingestion from ElasticSearch / Logstash

You can create a Logstash pipeline by stringing together plugins (inputs, outputs, and filters) in order to process data. To build a Logstash pipeline, create a config file to specify which plugins you want to use and the settings for each plugin.

A very basic pipeline might contain only an input and an output to forward all logs from ElasticSearch to Augtera Network AI Platform.

Input plugin can be configured to read from an Elasticsearch cluster, based on search query results and scheduled to run periodically according to a specific schedule e.g., every minute.

You can use two options as Output plugin in Logstash to forward logs to Augtera platform:

1. Syslog output plugin: this plugin will send events to Augtera as a syslog server. You can send messages compliant with RFC3164 or RFC5424 using UDP as the transport protocol. By default the contents of the message field will be shipped as the free-form message text part of the emitted syslog message.
2. Kafka ouput plugin: this plugin will write events to a Kafka topic. This output supports connecting to Kafka over SSL or Kerberos SASL. By default security is disabled but can be turned on as needed. The only required configuration is the topic_id. The full content of your events can be sent as JSON format. Logstash will encode your events with not only the message field but also with a timestamp and hostname.

Enabling AI/ML on Logs to Transform Network Operations

LogAI changes the log experience through high-performance, high-efficiency real-time AI/ML processing of streaming log messages from Syslog, or other non-standard format Logs such as Cloud Logs encoded as JSON and ingested using Kafka or Augtera APIs. The result is actionable and automated AI “needles” from the Log data haystack.

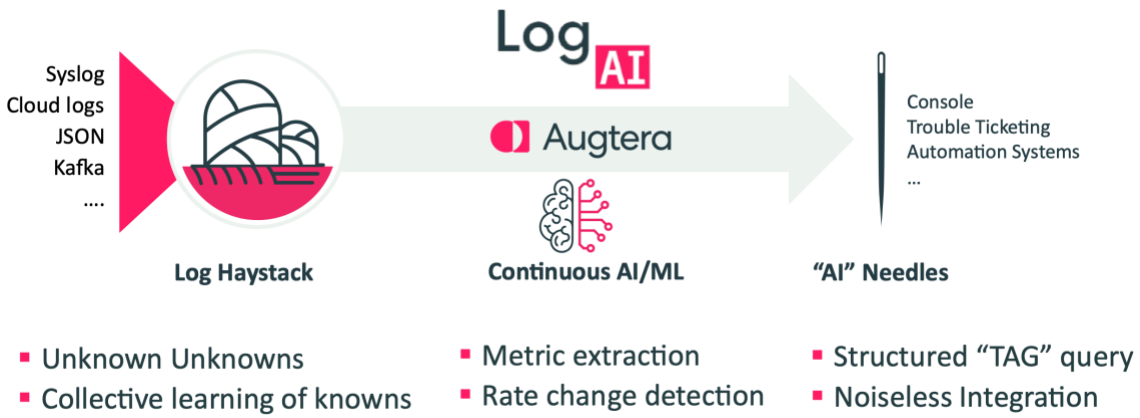
Network operations teams today either have no ability to process Logs in real-time or can only create limited rules to identify a set of known Log signatures. This leaves several gaps:

- Anomaly symptoms that have never been seen before i.e., the unknown unknowns
- Anomalies that are never acted on because they are lost in the noise
- Immediate implementation of new known Log signatures due to cumbersome rules-based approaches and/or lag time for internal software development and/or other inter-organization / process issues
- Extraction of metrics for anomaly detection & visualization
- Message rate-based anomaly detection with high fidelity and low noise

- Elimination of noise prior to signaling consoles, trouble ticket systems, and automation systems

Augtera Networks LogAI is a vastly different approach.

Image 3: LogAI Purpose-Built for Network Operations



Automated Workflows and Real-Time Notification

Simplified and flexible workflows for different teams driven by operator intent supporting ad-hoc analytics, notifications (slack, syslog, kafka), and automated ticketing integration (Service Now). Out of the box and custom Views with rich metadata aware filters (e.g., choose to notifiy only certain types of anomalies on certain types of devices) are used to define what types of anomalies, events and auto-correlated needles consoles, ticketing systems, and automation systems should be notified about.

Integration with third-party tools such as Splunk or ElasticSearch can be done using a syslog message directly sent by Augtera as a notification method.

It is also possible to send Augtera notifications to a kafka topics feeding Splunk or ElasticSearch.

- Splunk Connect for Kafka is a sink connector that allows a Splunk software administrator to subscribe to a Kafka topic and stream the data to the Splunk HTTP Event Collector.

After the Splunk platform indexes the events, you can then directly analyze the data or use it as a contextual data feed to correlate with other Kafka-related data in the Splunk platform.

- ElasticSearch can ingest Augtera notifications directly from a kafka topic using a Logstash pipeline similarly to the methodology described above, using kafla as input plugin and elasticsearch as output plugin.

Conclusion

Logs contain a wealth of information that operators increasingly want integrated into real-time workflows. In addition, Network Operations teams can no longer be simply reactive to incidents, they must proactively detect, see potential incidents before they occur, and prevent those incidents from ever happening.

The Augtera Network AI platform, including LogAI, was developed to not only enable network operations teams to react faster, but to reduce the need to react by eliminating noise and preventing future incidents. Further as described in this solution brief LogAI integrates with Splunk and ElasticSearch to enable seamless deployment in existing ecosystems.

Related Links

- LogAI Solution
- LogAI Solution Brief
- Augtera Data Sheet
- Configure Splunk syslog forwarder
- Send Splunk data to Kafka
- Splunk Connect for Kafka
- Configure a Logstash pipeline