Apache: Kafka

Apache Kafka is a distributed event store and stream-processing platform. It is an open-source system developed by the Apache Software Foundation written in Java and Scala. The project aims to provide a unified, high-throughput, low-latency platform for handling real-time data feeds.

Configuration

Guidance for onboarding data can be found in the Spunk Documentation, Getting Data In (Splunk Enterprise) or Getting Data In (Splunk Cloud). You might also find the following articles about Kafka configuration helpful:

- Integration reference: Kafka
- Splunk Data Stream Processor: Get data from Kafka
- Splunk Data Stream Processor: Send data to Kafka
- Splunk Connect for Kafka – Connecting Apache Kafka with Splunk

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. Built on top of the Kafka Connect library, this connector provides:

- High scalability, allowing linear scaling, limited only by the hardware supplied to the Kafka Connect environment.
- High reliability, by ensuring at-least-once delivery of data.
- Ease of data onboarding and simple configuration with Kafka Connect framework and Splunk's HTTP event collector.
- Support for ingestion of Kafka Record headers.

Data application

When your Splunk deployment is ingesting Apache Kafka data, you can use the data to achieve the following:

- Using Kafka to monitor at scale

You can also check this article for further guidance:

- About Splunk UBA Kafka ingestion in UBA and Splunk Cloud environments