Open Source Log Management Platform

SkaLogs

SENTAI is launching **SkaLogs**: a Centralized Log Collection Open Source Real Time Monitoring Big Data Platform

- SkaLogs is based on the ELK Stack (Elasticsearch, Logstash, Kibana), Kafka, and SkaETL.
- SkaETL is a dedicated ETL developed in Java by SkaLogs to handle both structured and unstructured Data.

Platform Features

SkaLogs is a self-hosted entreprise grade Open Source – Big Data – Real-Time Platform design to help you deploy, manage, and scale a centralized log management solution. It is based on 3 core components (Apache V2):

1. SkaETL : an Open Source (SkaETL GitHub Repo) real-time

ETL developed by SkaLogs,
2. ELK Stack : Elasticsearch, Logstash, Kibana,
3. Kafka.

Functional Architecture Diagram

Thanks to its advanced ETL (SkaETL), the deployed SkaLogs instance can be turned into several use-cases (Solutions):

- Centralized Log Management Platform,
- IT Operation Analytics (ITOA),
- Business Activity Monitoring (BAM)
- Security Information and Event Management (SIEM)

The SkaLogs Platform consists of a bundle which deploys many services, and scales them according to the ressources (cloud, on-premise) allocated to the instance deployed :

SkaLogs Bundle (GitHub repo)

The entire platform is deployed via a single Ansible script which:

- installs a bundle consisting of the above-mentioned 3 core components (SkaETL, ELK, Kafka),
- adds multiple side components,
- assembles the pieces into a scalable, automated, resilient, self-monitored, and complete end-to-end Log Management Platform,
- provides an entirely managed infrastructure (Rancher) with containerized (Docker) and orchestrated (Kubernetes) components.

The SkaLogs bundle includes:

- Rancher as a container management platform,
- SkaETL as an advanced log-dedicated ETL (SkaETL, developed by SkaLogs) with multiple guided workflows to help you with all the difficult tasks:

        - Logs: collect, transform, normalize, parse, aggregate,
        - Metrics: compute (before ES ingestion), store, search, investigate,
        - Alerts: create thresholds with alerts and notifications.
        - Visualize:
            - before ES ingestion: monitor data before ingestion and indexing in Elasticsearch
            - after ES ingestion:
                - Kibana as a front-end to Elasticsearch
                - Grafana as a front end for technical monitoring

Core Features

- Self-hosted (on-premise or cloud) complete end-to-end centralized Log Management Platform
- Scripted and Automated deployment (Ansible, Shell and yaml scripts)
- Container management (Rancher container management platform) and Orchestration (Kubernetes)
- Guided workflows for
    - data and Log ingestion and transformation (structured and unstructured)
    - real-time metrics computations and insights
    - interfacing with your own ML algorithms

SkaETL Features

SkaETL is a specialized 100% log-dedicated ETL developed by SkaLogs, allolwing you to process structured or unstructured data. The difficult task of data transformation is completely simplified thanks to multiple guided workflows:

- Ingest, parse, transform, enrich, normalize, aggregate,

index, archive
- Compute (simple statistics, complex functions, and ML algorithms)
- Search and investigate
- Visualize and monitor
- Alerts and notifications

Technical Features

- Microservices based architecture
- Packages multiple Open Source Libraries and framework
- Entirely managed infrastructure with containerized and orchestrated components
- Base deployment assembles 50+ services into 150+ docker containers
- Scalable, automated, resilient, self-monitored
- Error retry mechanism
- Volume: Scales without limits
- Speed: ingest at +100 K EPS (events / second or json documents / second)

Technical Features List

Open Source – Big Data – Real time (Kafka®) – Advanced ETL (SkaLogs ETL) – {Indexing, Storage, Archiving} (Elasticsearch®, HDFS) – Self-hosted (On-premise, Cloud) – {private, public, hybrid, multi}-Cloud – Multi-{cluster, instance, tenant} Environments – Virtualization (OpenStack, KVM) – Containerization (Docker) – Orchestration (Kubernetes) – Managed Infrastructure (Rancher) – Scripted (Ansible, Shell, yaml) – Automated Deployment (Ansible) – Secured (SSL, Elastic Security, IPsec) – Managed Updates – Directory Connectors (LDAP, Kerberos, AD) – Self-Monitored – Error Retry Mechanism – Storage Connectors – Metrics and Computations (Statistical, Advanced functions, Machine Learning) – Visualization Templates (Grafana, Kibana) –

Notifications (thresholds, SMS, email, Slack…)