

Data Analytics

Data Flow



다양한 소스로부터 데이터를 추출/변환/전송하고 데이터 처리 흐름을 자동화하는 서비스

Data Flow는 다양한 데이터 소스로부터 대용량의 데이터를 추출하고, 스트림/배치 데이터의 변환/전송에 대한 처리 흐름을 시각적으로 작성하는 데이터 처리 흐름 도구로, 오픈소스 Apache NiFi를 제공합니다. Data Flow는 Samsung Cloud Platform의 Kubernetes Engine 클러스터 환경에서 단독으로 사용하거나, 다른 애플리케이션 SW와 함께 사용할 수 있습니다.

편리한 설치

Data Flow는 표준 Kubernetes 클러스터 환경에서 웹 기반 콘솔을 통해 손쉽게 설치가 가능합니다. 오픈소스 Apache NiFi 기반의 확장형 클러스터링에 필요한 아키텍처를 자동으로 구성하여, ZooKeeper, Registry 및 관리 모듈이 자동 설치됩니다. Data Flow를 통해 서비스 연결에 필요한 설정 파일, NiFi 템플릿 등을 설정/배포할 수 있습니다.

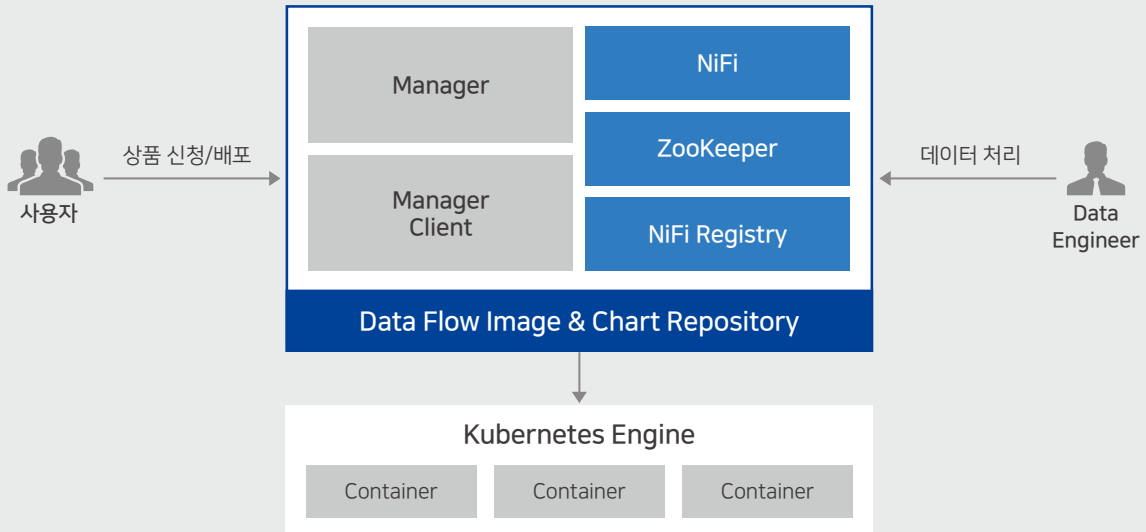
손쉬운 데이터 흐름 관리

스트림/배치 데이터의 처리 흐름을 사용자 환경에 맞게 GUI 기반으로 쉽게 작성할 수 있으며, GUI 기반의 데이터 처리 흐름 작성으로 시스템 간 데이터를 효율적으로 추출/전송/처리할 수 있습니다.

NiFi 템플릿 갤러리

레퍼런스 NiFi 템플릿을 공유/배포 할 수 있습니다. Data Flow는 현업에서 자주 사용하는 데이터 처리 흐름에 대한 작업 파일을 갤러리로 제공하며, 사용자는 자신이 작성한 데이터 처리 흐름 작업을 공유할 수 있습니다.

서비스 구성도



주요 기능




- 간편한 설치
 - 웹 콘솔을 통해 컨테이너 기반의 오픈 소스 NiFi 확장형 클러스터 환경 자동 설치
 - 초기 설치 항목 : NiFi, ZooKeeper, NiFi Registry
- 데이터 처리
 - 연결 가능 데이터 소스 : HDFS, HIVE, HBASE, Solr, Cassandra, MongoDB, Elasticsearch, Kafka, RabbitMQ, Syslog, HTTPS, SFTP 등
 - 실시간 데이터 연결, 데이터의 이동 경로 추적, 데이터 분실 방지
 - 대용량 데이터 및 병렬 작업 처리
- GUI 기반의 편리한 관리
 - GUI 기반으로 NiFi 설정 및 Nar 파일에 대해 클러스터 환경에서 배포/관리
 - NiFi 템플릿 공유/배포 기능

요금 기준

- 과금
 - Data Flow가 사용하는 Kubernetes Engine의 Pod들의 CPU 사용시간

FOR MORE INFORMATION

SAMSUNG SDS

 www.samsungsds.com / cloud.samsungsds.com
 contact.sds@samsung.com / scp_sales@samsung.com
 youtube.com/samsungsds

