

Apache NiFi es un proyecto de código abierto desarrollado por la Apache Software Foundation que proporciona una plataforma escalable y extensible para automatizar el flujo de datos, la integración de sistemas y la gestión de datos en tiempo real. NiFi es especialmente útil en entornos empresariales y de Big Data donde la ingesta, el procesamiento y la transferencia de datos son fundamentales.

Algunas de las características clave de Apache NiFi incluyen:

- Ingesta de datos: NiFi puede recolectar datos de una variedad de fuentes, como archivos locales, sistemas de mensajería, bases de datos, servicios web, sensores y más.
- Procesamiento de datos en tiempo real: Permite realizar transformaciones y enriquecimientos en tiempo real en los datos mientras fluyen a través del sistema. Esto es útil para la limpieza de datos, la agregación, la validación y otras operaciones.
- Enrutamiento y enriquecimiento: NiFi permite enrutar datos en función de reglas predefinidas y enriquecerlos con metadatos adicionales o información contextual.
- Seguridad: Ofrece características de seguridad robustas, como autenticación, autorización y cifrado de datos, para proteger la información confidencial.
- Escalabilidad: NiFi es altamente escalable y puede manejar grandes volúmenes de datos y flujos de trabajo complejos.
- Interfaz de usuario intuitiva: Viene con una interfaz gráfica de usuario que facilita la configuración y la supervisión de los flujos de datos.
- Extensibilidad: NiFi es extensible a través de extensiones personalizadas y permite la integración con otras herramientas y servicios.
- Monitorización y supervisión: Proporciona capacidades de monitorización y registro detallado para rastrear el rendimiento y la salud de los flujos de datos.

Nota: Apache NiFi se utiliza en una variedad de casos de uso, como la ingesta de datos en sistemas de Big Data, la automatización de flujos de trabajo de ETL (extracción, transformación y carga), la integración de sistemas empresariales, la recolección de datos de sensores en el Internet de las cosas (IoT) y más. En resumen, es una herramienta versátil para gestionar el flujo de datos en una organización y simplificar tareas de integración y procesamiento de datos.

Referencia

- Documentación & Videos <https://nifi.apache.org/videos.html>
- Contenedor: <https://hub.docker.com/r/apache/nifi>