

APACHE NIFI

Aplicado a la AO



¿QUÉ ES APACHE NIFI?

→ Es una herramienta open-source de automatización de flujos de datos diseñada para la recopilación, procesamiento y distribución de datos en tiempo real y por lotes. ←

NOS PERMITE

Envío de eventos a sistemas de streaming

Ingesta en distintas bases de datos

Interfaz gráfica de fácil uso

Procesamiento de datos

ARQUITECTURA

FlowFiles

Son las unidades básicas de datos que se procesan en la herramienta

Cada una de estas unidades consta de:

- Contenido
- Atributos

Repositorios

Parte fundamental de la herramienta y cuenta con 3 tipos diferentes

- FlowFile Repository
- Content Repository
- Provenance Repository

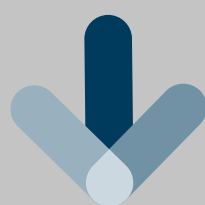
Controladores

Configuraciones comparidas que pueden ser utilizadas por multiples procesadores

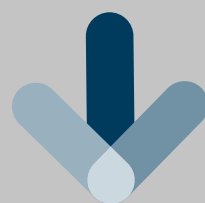
Procesadores como los que permiten la ingesta en bases de datos, almacenamiento en caché, etc

CLÚSTER Y ALTA DISPONIBILIDAD

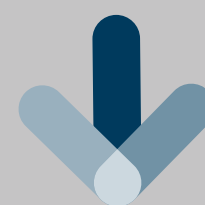
Nifi puede ejecutarse en modo clúster, donde múltiples nodos trabajan juntos para procesar grandes volúmenes de datos.



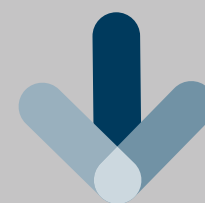
**PRIMARY
NODE**



Nodo responsable
de ejecutar ciertos
procesadores
configurados



**COORDINATOR
NODE**



Nodo encargado de
la gestión y
orquestración del
clúster

ZOOKEEPER →

Es el encargado de coordinar la comunicación entre nodos y mantener el estado del sistema.

CASOS DE USO

1

**Integración de
múltiples fuentes
de datos**

2

**Procesamiento
en tiempo real**

3

**Carga de datos
en BBDD y
almacenamiento
en la nube**

4

**Automatización
de flujos de datos**