

# NiFi Notes

---



<https://nifi.apache.org/docs/nifi-docs/html/overview.html>

Johnathan  
Brummett

---

---

---

A small, dark grey icon of a pen nib pointing upwards and to the right, located at the end of the third horizontal line.

Nifi è un software per automatizzare e gestire i flow tra i sistemi. Un sistema affidabile e potente per processare dati.

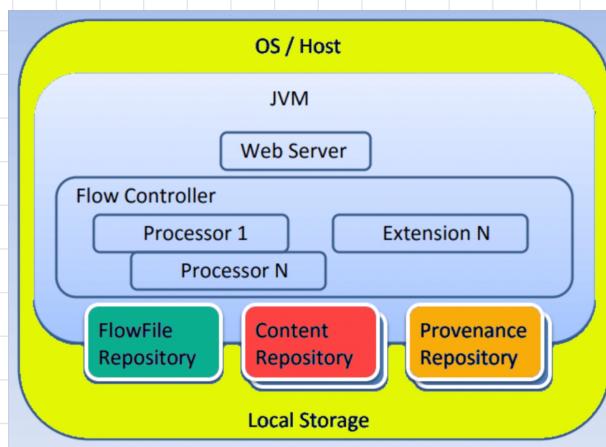
Fornisce una GUI web per creare, monitorare e gestire i flussi. Nifi è facilmente estendibile con lo sviluppo di componenti custom.

Nifi permette di:

- Permette di caricare dati da numerosi sorgenti
- Controllo redirezione per gestire lo spostamento dei dati tra sorgente e destinazione
- Visualizzazione enterprise
- Permette di usare librerie esistenti
- Facile da integrare
- Design adatto a scalare nell'cluster
- Visualizzare e monitorare le performance
- Aiuta a lanciare e stoppare Job a livello di gruppo
- Interfaccia Drag & Drop per implementare i flussi.

Apache Nifi ha un'architettura basata su flowfile, una volta che i dati vengono presi da sorgenti esterne vengono rappresen-

ta in nell'architettura Nifi sotto Flowfile



**FLOW FILE** = Dati originali con meta-information. Permette di processare ogni tipo di dato

**FLOW FILE PROCESSOR** = Permette di operazioni in blocco

**FLOW CONTROLLER** = Tieni un record su come i processi sono connessi, gestisce thread e allocazioni tra i vari processi.

**WEB SERVER** = Host per HTTP e API

**EXTENSION** = C'è sono molti tipi di extension Nifi, che operano nella JVM

**CONNECTION** = Un link tra i processi che comunicano

**BACK PRESSURE** = Ferma il sistema dell'output in stress controllando la quantità dei dati nelle code

**PROCESS GROUP** = Un insieme di processi e le loro connessioni che comunicano.

**FLOWFILE REPOSITORY** = Tieni traccia dello stato dei dettagli dei flowfile all'interno del flow

**CONTENT REPOSITORY** = Un area dove il contenuto viene salvato in blocchi di un Flowfile

**PROVENANCE REPOSITORY** = Un area dove vengono raccolti tutti i dati sugli eventi di provenienza.

### Apache NiFi Features

- NiFi supports buffering of all queued data and offers an ability of back pressure as those queues may reach specified limits
- NiFi allows the setting of one or more prioritization schemes
- Provides connection processors for many data sources
- Supports any device which runs Java
- Ideal for limited connectivity places
- Support for troubleshooting and flow optimization
- Offers role-based authentication/authorization
- Allows download, recovery, and replay of individual files
- Build your processors, controller services, and more
- Provide content encryption, communication over secure protocols
- Enables rapid development and effective testing
- Allows for the development of simple single-function components that can be reused and combined to make more complex flows
- Allows classloader isolation for easier management of dependencies

### Best practices Running Apache NiFi

- Ideal to separate test/dev/production environments in NiFi
- You should break your flow into process groups
- Use a naming convention, use comments and labels
- Organize your projects into three parts ingestion, test & monitoring
- Use unique names for variable

### Disadvantage of Nifi

- Need precise security and compliance controls
- You need to know the underlying system very well while working with Apache NiFi
- Must maintain chain of custody for data
- Transport / Messaging may not prove enough
- Data access needs exceed available resources to transport
- Not all data is created equally
- SSL and topic level authorization may not be sufficient