

Nifi Übung 1: UI Elemente

Schritt 1: Strukturiertes Arbeiten:



Erschaffen einer Process Group (Drag-and-drop in weiß kariertes Feld):

Geben Sie einen Namen bei der Erschaffung an, z.B. "Orders".

Durch Doppelklick auf die Process Group können sie diese öffnen.

Process Groups:

Strukturelles Element zur Ordnung von Flow Prozessen. Ähnlichkeit zur Ordnerstruktur in z.B. Windows. Dadurch wird ein dedizierter Workspace für z. B. ein Projekt geschaffen.

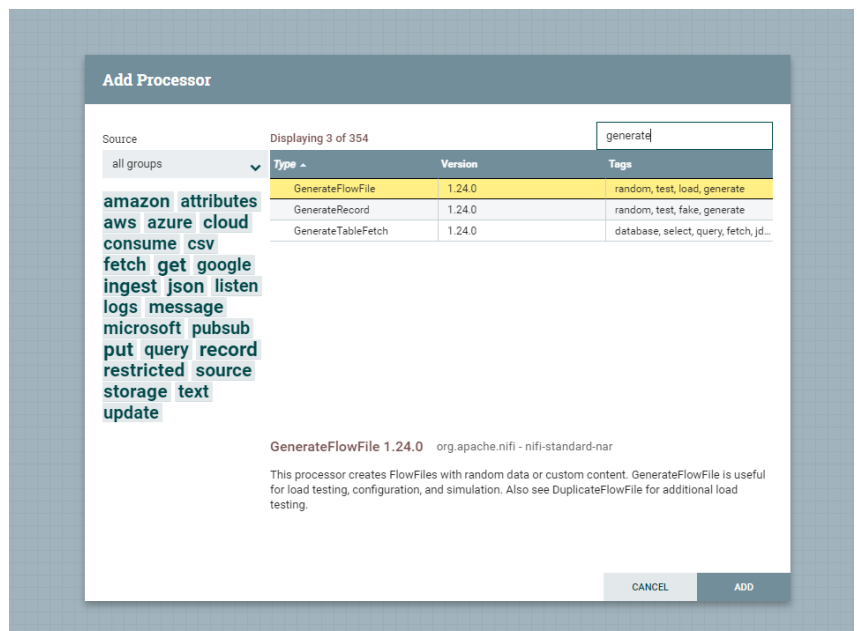
Schritt 2: Start eines Flow's:



Erschaffen eines Processors (Drag-and-drop in weiß kariertes Feld):

Zur Generierung von Daten zu Testzwecken benötigen wir einen Processor des Typs "**Generate Flow File**". Dazu in der UI nach dem Drag-and-drop des Processor Icons nach diesem Typ suchen und auswählen. Zuletzt die Auswahl durch "**ADD**" bestätigen.

Diesen Schritt wiederholen, um einen Processor des Typs "**ConvertRecord**" zu erschaffen.

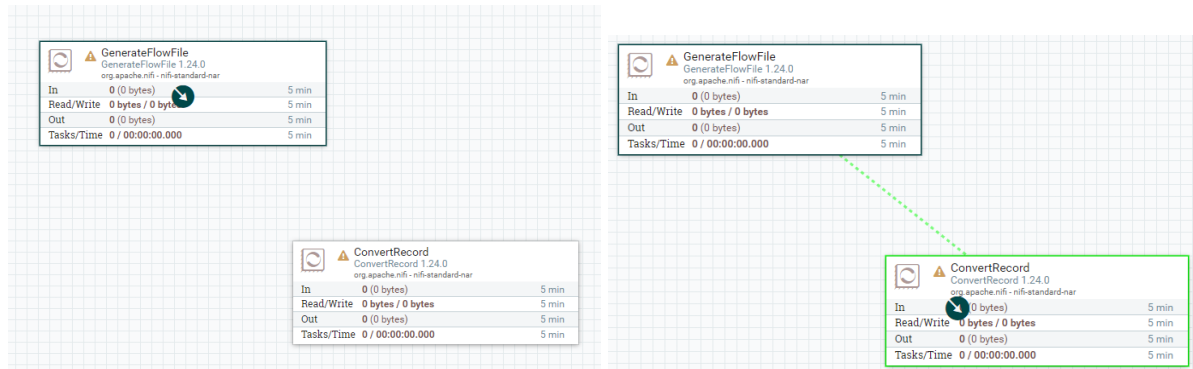


Processor:

Hauptelement um etwas mit Daten zu machen, ob Extraction, Transformation, Augmentation oder Filtern;

Die Prozessoren verbinden:

Über Auswählen des Pfeil-Icons und Drag-and-drop auf den anderen Processor:



Auswahl der Relationship:

“Success” (Einzig verfügbare Relationship bei dem Generate Processor, jedoch bei anderen Processors sind auch andere möglich. Dabei wird das weitere Verfahren gemappt; beispielsweise, falls die Verarbeitung fehlgeschlagen ist, muss mit dem FlowFile anders verfahren werden und dieses muss einen anderen Flow nehmen, als bereits verarbeitbare FlowFiles.)

Finales bestätigen mit **“ADD”**

Flow starten:

Es kann durch rechtsklick auf GenerateFlowFile der Prozess gestartet werden.

Dazu sind die Optionen **“Start”** und **“Run Once”** verfügbar.

“Start”:

Führt den processor nach eingestelltem scheduling aus (CRON, Timer based)

“Run Once”:

Lässt den Processor ein Element bearbeiten.

Konfiguration des scheduling:

Doppelklick auf Kachel (des Processors, den man einstellen möchte) →

SCHEDULING → Scheduling Strategy auswählen usw.

Beobachtung:

Die FlowFiles bleiben auf dem Weg zu ConvertRecord stecken (weil ConvertRecord diese nicht verarbeitet).

Mit Rechtsklick auf die Kachel des Pfeils und Auswahl von **“List queue”** können die feststeckenden FlowFiles besichtigt werden.

Nifi Übung 2: Konfiguration eines Flows

1. Konfiguration der Processors im Flow:

Generation Processor einstellen, um ein geordnetes JSON zu erzeugen:
Doppelklick auf “**GenerateFlowFile**” Kachel → PROPERTIES → Custom Text:
einfügen:

```
{  
  "name": "Batman"  
}
```

Dann noch bei “**Mime**
Type” eintragen:

application/json

Durch Apply bestätigen

Property	Value
File Size	0B
Batch Size	1
Data Format	Text
Unique FlowFiles	false
Custom Text	{...}
Character Set	UTF-8
Mime Type	application/json

ConvertRecord Processor einstellen:

1. Doppelklick auf “**ConvertRecord**” Kachel → PROPERTIES → Record Reader → Reference parameter → wähle “Create new service”
2. Compatible Controller Services → JsonTreeReader
3. Controller Service Name: passenden Namen vergeben
4. Durch “**Create**” bestätigen
5. Selbiges mit Writer machen in einem Format ihrer Wahl (Vorschlag: CSV)
6. Aktivieren der Services: auf einen der beiden Pfeile rechts in dem Fenster drücken → im darauffolgenden Fenster bei beiden Services auf den Blitz klicken, um diese zu aktivieren.

Der “**Convert Record**” Processor ist noch im Modus “Invalid”, um das zu ändern, muss dieser Processor einen Output bei den Relationen “**Failure**” und “**Success**” besitzen.

Dazu zwei neue Processors erzeugen (siehe Übung 1):

“**PutEmail**” kann z.B. bei relationship **failure** verwendet werden, um Alarm zu schlagen

“**UpdateAttribute**” falls das Schema wie erwartet konvertiert wird, soll ein Attribut geändert werden (symbolisch für mögliche Weiterverarbeitung bzw. Verarbeitbarkeit)

Diese zwei Prozessoren als Output Ziele von **ConvertRecord** auf die jeweiligen Relationships setzen.

Die Konfiguration der gerade erstellten Knoten ist nicht Teil der Übung und kann ignoriert werden.

Als letzten Punkt die Prozessoren "**GenerateFlowFile**" und "**ConvertRecord**" laufen lassen (siehe Übung 1). Generierte FlowFiles sollten in der Relationship "**success**" vor dem Processor "**UpdateAttribute**" stecken bleiben.

Wie man die Musterlösung öffnet:

erstellen einer Process Group:



- Icon in karierte Fläche ziehen

im neuen Fenster auf das Zeichen rechts klicken:

Laden der Flow Definition im Json Format:

ExampleFlow.json

Leider gibt es ein Problem, dass die Reader und Writer Services noch aktiviert werden müssen:

Doppelklick auf **ConvertRecord** Knoten/ Processor → PROPERTIES Reiter wählen → auf einen der beiden Pfeile rechts in dem Fenster drücken → im darauffolgenden Fenster bei beiden Services auf den Blitz klicken, um diese zu aktivieren.