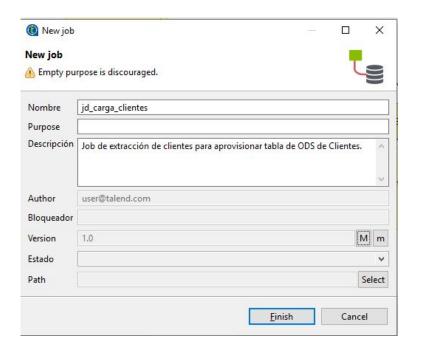
DW y ETL TALEND OPEN STUDIO JOBS



ÍNDICE

- 1. CREACIÓN DE JOBS
- 2. INCORPORAR COMPONENTES
- 3. SUBJOBS
- 4. CONFIGURACIÓN VISTA EDICIÓN
- 5. VISTA RUN DEL JOB
- 6. TIPOS DE FLUJO
- 7. ORGANIZACIÓN DE JOBS
- 8. PREJOB Y POSTJOB

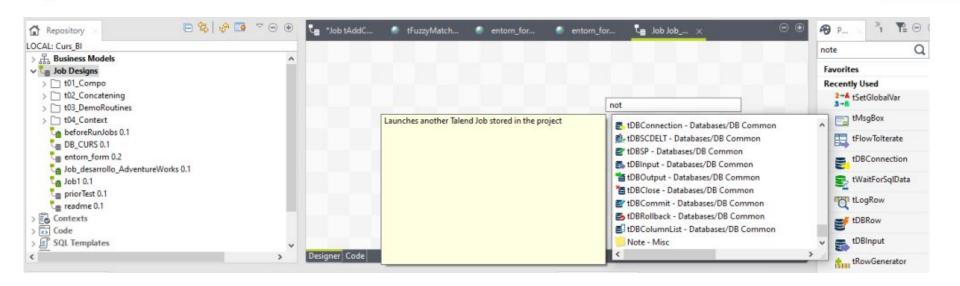
CREACIÓN DE JOBS



Un job de Talend es un proceso que vamos a desarrollar a partir de diferentes componentes conectados mediante flujos de datos y disparadores.

La creación de un job vuelve a ser similar a lo que ya hemos tratado en los otros elementos del repositorio. Empezamos con un click en el botón derecho en el nodo de **Job Designs** y posteriormente seleccionamos **Create job**. O **Create folder** si queremos organizar los jobs por carpetas.

INCORPORAR COMPONENTES



Para incorporar componentes, podemos usar la Paleta situada a la derecha de la pantalla, también podemos usar el curso, escribiendo directamente el nombre del componente, y también podemos usar elementos del repositorio.

SUBJOBS



Cuando unimos componentes y conexiones formamos diferentes **subjobs** que identificaremos por estar contenidos en una forma rectangular que los rodean que por defecto es de color azul.

El componente principal si no tiene dependencia subordinada de otros subjobs, nos podemos fijar que está resaltado en color verde.

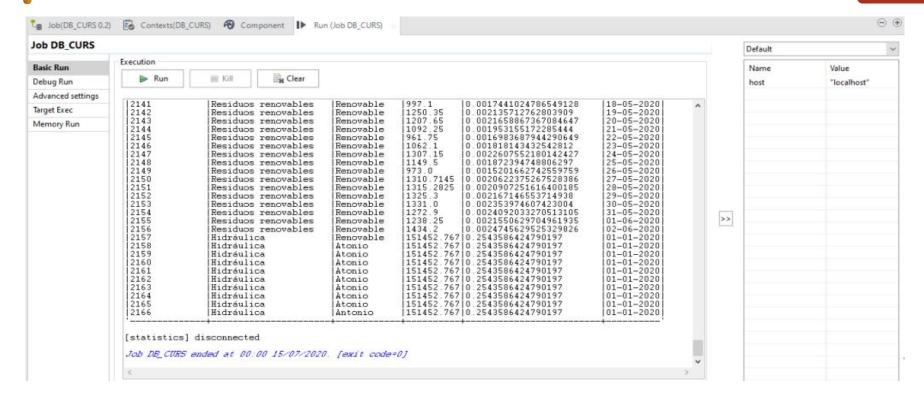
CONFIGURACIÓN VISTA EDICIÓN

La vista de edición corresponde a la vista del panel inferior central de la pantalla y dispone de 4 pestañas:

- **Job**: Muestra los datos generales del job que hemos introducido.
- Contexts: Sirve para incorporar parámetros que se utilizarán como variables de contexto que nos permite establecer diferentes escenarios de ejecución del job.
- Component: Muestra y permite configurar los diferentes elementos del componente que tengamos seleccionado. Si seleccionamos un componente del lienzo, esta pestaña mostrará los datos de configuración concretos del componente seleccionado. Por otra parte si en vez de seleccionar el componente, seleccionamos el flujo que conecta dos componentes, esta misma pestaña nos mostrará información del esquema de datos. Por último si seleccionamos el subjob, mediante un clic en el área del subjob, podemos configurar elementos del aspecto del subjob, como el color del subjobs
- Run: Desde esta vista disponemos de botones desde los cuales podemos ejecutar el job, pararlo «kill» o limpiar la pantalla del job. También nos muestra el log de ejecución o los datos de una query. En la derecha una tabla muestra las variables de contexto y contenido de las mismas.



VISTA RUN DEL JOB



TIPOS DE FLUJO

Flujo de datos Main

Flujo que hace circular todo el caudal de datos principal del job.

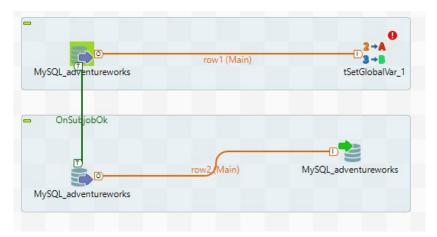
Flujo de datos Iterate

Un flujo de datos tipo Iterate permite enviar los registros de uno a uno y son captados por una variable que podemos utilizar en un componente de manera iterativa. Como por ejemplo para enviar correos electrónicos a partir de los registros de una tabla de datos

Disparadores Triggers

Los disparadores son otro tipo de conectores que nos permiten activar diferentes caminos de subjobs en función del resultado OK/Error de un subjob anterior.

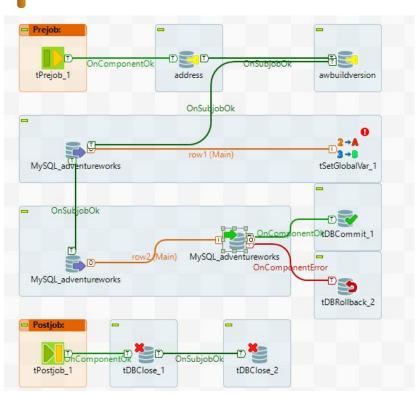
ORGANIZACIÓN DE JOBS



Para asegurar el orden de ejecución de los jobs es necesario el uso de **disparadores o triggers**.

Como nos muestra la imagen siguiente, el primer job se convierte en subjob principal de ambos subjobs. El subjob segundo pierde el color verde del componente principal.

PREJOB Y POSTJOB



Los componentes **prejob** y **postjob** permiten realizar operaciones antes y después de la ejecución de un job.

Tareas de prejob:

- Cargar información requerida por un subjob.
- Comprobar la existencia de un fichero.

Tareas de postjob:

- Limpieza de ficheros temporales.
- Ejecutar una tarea, aunque el job haya fallado.



Madrid Barcelona Bogotá

Jon Pineda Galvez

jpinedagalvez@gmail.com
https://www.linkedin.com/in/jonathanpg