Contents

[Updating attributes and writing files 1](#_Toc202965937)

[What is the purpose of the ConvertAvroToJSON processor? 1](#_Toc202965938)

[Why do we need to convert Avro to JSON before converting to CSV? 2](#_Toc202965939)

[Updating attributes and writing files 2](#_Toc202965940)

[What is the objective of the data flow in this video? 2](#_Toc202965941)

[What processor is used to update the filename attribute? 2](#_Toc202965942)

## Add GetFile and PutDatabase processors

* **Setting Up GetFile Processor**: The GetFile processor is configured to read files from a specified directory in the local file system and delete them after reading to avoid duplication.
* **Configuring Run Schedule**: The Run Schedule for the GetFile processor is set to check for new files every five seconds.
* **Using PutDatabaseRecord Processor**: The PutDatabaseRecord processor is used to write records from the FlowFile into a PostgreSQL database table, supporting various SQL servers like MySQL and PostgreSQL.

These steps help in automating the process of reading data from files and storing it in a database using Apache NiFi.

## Configure processors to convert data formats

* **Conversion Process**: The video explains how to convert data from Avro format to JSON format using the ConvertAvroToJSON processor, and then from JSON to CSV format using the ConvertRecord processor.
* **Configuration**: It details the configuration of the Record Reader and Record Writer controller services, specifically setting up JsonTreeReader for reading JSON records and CSVRecordSetWriter for writing CSV records.
* **Connections and Relationships**: The video shows how to connect the processors and configure relationships to ensure the data flow processes correctly, including setting up termination on failure.

These steps ensure that the data is successfully converted and written out in the desired format.

A screenshot of a computer

AI-generated content may be incorrect.

### What is the purpose of the ConvertAvroToJSON processor?

The purpose of the ConvertAvroToJSON processor is to convert data from the Avro format to the JSON format. This is necessary because there is no direct way to convert Avro to CSV. By first converting Avro to JSON, you can then use another processor to convert JSON to CSV, allowing the data to be written out as a CSV file on your local machine.

### Why do we need to convert Avro to JSON before converting to CSV?

We need to convert Avro to JSON before converting to CSV because there is no direct way to convert data from Avro format to CSV format. The intermediate step of converting Avro to JSON allows us to then use another processor to convert JSON to CSV. This two-step process ensures that the data can be properly formatted and written out as a CSV file on your local machine.

## Updating attributes and writing files

* **UpdateAttribute Processor**: This processor updates the attributes of a FlowFile, such as the filename, using the Attribute Expression Language.
* **PutFile Processor**: This processor writes the FlowFile to the local file system, using the updated filename attribute.
* **Configuration**: Ensure the Record Reader and Record Writer Controller Services are enabled for the ConvertRecord processor to execute the data flow successfully.

### What is the objective of the data flow in this video?

The objective of the data flow in this video is to write out grouped and aggregated data as a CSV file on the local file system. This involves updating the filename attribute using the UpdateAttribute processor and then using the PutFile processor to write the final FlowFile to the local system.

### What processor is used to update the filename attribute?

The processor used to update the filename attribute in the data flow is the **UpdateAttribute** processor. This processor uses the Attribute Expression Language to update the attributes associated with the FlowFile.

A screenshot of a computer

AI-generated content may be incorrect.