

# MYSQL to Oracle Flowfiles Template using NIFI:

## Steps Followed:

1. Used Generate Table Fetch processor and configured partition size to 30000 which is the maximum number of records that is loaded into the memory at once.

### Configure Processor

SETTINGS	SCHEDULING	PROPERTIES	COMMENTS
----------	------------	------------	----------

Required field +

Property	Value
Database Connection Pooling Service	MySQLConnect →
Database Type	MySQL
Table Name	cx_event_log
Columns to Return	No value set
Maximum-value Columns	No value set
Max Wait Time	0 seconds
Partition Size	30000
Column for Value Partitioning	No value set
Additional WHERE clause	No value set
Output Empty FlowFile on Zero Results	false

CANCEL

APPLY

Cannot use QueryDatabaseTable as it fetches all the records into the memory before separating it into the flowfiles.

2. Used ExecuteSql processor to execute the sql queries from the previous processor to produce the flowfiles with the records.

Configuration:

Max Rows per flow file = 4

OutputBatchSize = 5 # To increase efficiency chose small numbers.

## Configure Processor

[SETTINGS](#)[SCHEDULING](#)[PROPERTIES](#)[COMMENTS](#)

Required field

Property		Value	
Database Connection Pooling Service	?	MysqlConnect	→
SQL Pre-Query	?	No value set	
SQL select query	?	No value set	
SQL Post-Query	?	No value set	
Max Wait Time	?	0 seconds	
Normalize Table/Column Names	?	true	
Use Avro Logical Types	?	false	
Compression Format	?	NONE	
Default Decimal Precision	?	10	
Default Decimal Scale	?	0	
Max Rows Per Flow File	?	4	
Output Batch Size	?	1000	

[CANCEL](#)[APPLY](#)

The following conversions were done to avoid timestamp errors. On direct input of the records obtained from the previous processor into the putDatabaseRecord Processor the date read from the MYSQL database was not being identified as date and hence would produce a date read error on inserting to ORACLE.

Tried accepting as string and converting to date, which worked, but reduced the efficiency and increased the time of the entire conversion, hence went with avro to json conversion which accurately retained the timestamp and date attributes.

- Used ConvertRecord processor to read the records in Avro Format and convert to JSON format to preserve the Attribute Types(METADATA).

## Configure Processor

SETTINGS

SCHEDULING

PROPERTIES

COMMENTS

Required field



Property		Value	
Record Reader	?	AvroReader	→
Record Writer	?	JsonRecordSetWriter	→
Include Zero Record FlowFiles	?	true	

CANCEL

APPLY

5. Since the JSON records included the attribute type and also the record converted the JSON records to sql insert statements using CONVERTJSONTOSQL processor.

## Configure Processor

SETTINGS

SCHEDULING

PROPERTIES

COMMENTS

Required field



Property		Value	
JDBC Connection Pool	?	Oracleconnect	→
Statement Type	?	INSERT	
Table Name	?	CX_EVENT_LOG	
Catalog Name	?	No value set	
Schema Name	?	No value set	
Translate Field Names	?	true	
Unmatched Field Behavior	?	Ignore Unmatched Fields	
Unmatched Column Behavior	?	Fail on Unmatched Columns	
Update Keys	?	No value set	
Quote Column Identifiers	?	false	
Quote Table Identifiers	?	false	
SQL Parameter Attribute Prefix	?	sql	
Table Schema Cache Size	?	100	

CANCEL

APPLY

7. Since the SQL INSERT queries had been generated, hence used PUTSQL processor to execute it into the ORACLE database.

Configure Processor

SETTINGS

SCHEDULING

PROPERTIES

COMMENTS

Required field

Property		Value	
JDBC Connection Pool	?	Oracleconnect	→
SQL Statement	?	No value set	
Support Fragmented Transactions	?	true	
Database Session AutoCommit	?	false	
Transaction Timeout	?	No value set	
Batch Size	?	100	
Obtain Generated Keys	?	false	
Rollback On Failure	?	false	

CANCEL

APPLY

Didn't use PUTDATABASERECORD processor as it would cause the timestamp error again.

CONTROLLERSERVICES:

DBtoCSV Configuration

GENERAL

CONTROLLER SERVICES

Name	Type	Bundle	State	Scope	
AvroReader	AvroReader 1.9.2	org.apache.nifi - nifi-record-serialization-ser...	Enabled	DBtoCSV	
JsonRecordSetWriter	JsonRecordSetWriter 1.9.2	org.apache.nifi - nifi-record-serialization-ser...	Enabled	DBtoCSV	
MysqlConnect	DBCPCConnectionPool 1.9.2	org.apache.nifi - nifi-dbcpc-service-nar	Enabled	DBtoCSV	
Oracleconnect	DBCPCConnectionPool 1.9.2	org.apache.nifi - nifi-dbcpc-service-nar	Enabled	DBtoCSV	

Last updated: 16:03:13 IST

Listed services are available to all descendant Processors and services of this Process Group.

1. Avro Reader Configuration: Default








## Controller Service Details

SETTINGS

PROPERTIES

COMMENTS

### Required field

Property	Value	
<b>Schema Access Strategy</b>	 <b>Use Embedded Avro Schema</b>	
Schema Registry	 No value set	
Schema Name	 \${schema.name}	
Schema Version	 No value set	
Schema Branch	 No value set	
Schema Text	 \${avro.schema}	
<b>Cache Size</b>	 <b>1000</b>	

OK

## 2. JSON record set writer Configuration: Default:















## Controller Service Details

SETTINGS

PROPERTIES

COMMENTS

### Required field

Property	Value	
<b>Schema Write Strategy</b>	 <b>Do Not Write Schema</b>	
Schema Cache	 No value set	
<b>Schema Access Strategy</b>	 <b>Inherit Record Schema</b>	
Schema Registry	 No value set	
Schema Name	 \${schema.name}	
Schema Version	 No value set	
Schema Branch	 No value set	
Schema Text	 \${avro.schema}	
Date Format	 No value set	
Time Format	 No value set	
Timestamp Format	 No value set	
<b>Pretty Print JSON</b>	 <b>false</b>	
<b>Suppress Null Values</b>	 <b>Never Suppress</b>	
<b>Output Grouping</b>	 <b>Array</b>	

OK

### 3. DBCONNECTIONPOOL service : MYSQL CONNECTION:

## Controller Service Details

SETTINGS

PROPERTIES

COMMENTS

Required field

Property	Value
Database Connection URL	jdbc:mysql://engproddb.customerxps.com/clari...
Database Driver Class Name	com.mysql.jdbc.Driver
Database Driver Location(s)	file:///home/roshan/nifi-1.9.2/bin/mysql-connect...
Kerberos Credentials Service	No value set
Database User	clari5ai_efm
Password	Sensitive value set
Max Wait Time	500 millis
Max Total Connections	8
Validation query	No value set
Minimum Idle Connections	0
Max Idle Connections	8
Max Connection Lifetime	-1
Time Between Eviction Runs	-1
Minimum Evictable Idle Time	30 mins

OK

### 4. DBCONNECTIONPOOL service : ORACLE CONNECTION:

# Controller Service Details

SETTINGS

PROPERTIES

COMMENTS

Required field

Property		Value
Database Connection URL	?	jdbc:oracle:thin:@192.168.5.70:1521:db12C
Database Driver Class Name	?	oracle.jdbc.driver.OracleDriver
Database Driver Location(s)	?	file:///home/roshan/Downloads/dbeaver/ojdbc8...
Kerberos Credentials Service	?	No value set
Database User	?	cxpsadm_swaraj_48dev
Password	?	Sensitive value set
Max Wait Time	?	500 millis
Max Total Connections	?	8
Validation query	?	No value set
Minimum Idle Connections	?	0
Max Idle Connections	?	8
Max Connection Lifetime	?	-1
Time Between Eviction Runs	?	-1
Minimum Evictable Idle Time	?	30 mins

OK