

Kafka Connect FAQ

How do I change the output data format of a SinkConnector?

The format that is written to the output system is dependent on the SinkConnector itself. The SinkConnector will translate records into the output data format from the format produced by the converter specified in `key.converter` and `value.converter`. Please refer to the documentation for the specific connector to see supported output formats.

Why does a connector configuration update trigger a task rebalance?

Not all connector configuration updates will trigger a task rebalance, but most will. This is because the trigger for the task rebalance is actually a need to reconfigure tasks in a way that that Connect cannot safely do without a rebalance. There are two basic reasons for a rebalance to happen:

1. The total number of connectors or tasks has changed, requiring reassignments to happen within the active workers.
 2. Configuration of tasks change in a way that Connect cannot be sure if tasks will need to coordinate to ensure correct behavior. For example, we want to avoid skipping or repeating messages if partitions are reassigned.
-

Why should I use distributed mode instead of standalone?

We recommend using distributed mode for most production use cases. Please see [here](#) for details.

Do I need to write custom code to use Kafka Connect?

The [Connector Hub](#) has many Connectors that cover many use cases such that most users will not need to write any code. If your use case is not covered, it may require you to extend one of these Connectors or write a new one. If you find that to be the case, please see our [developer guide](#).

Is Schema Registry a required service to run Kafka Connect?

No, it is not required, but it is recommended if you plan to use Avro for a data format. This is because it can help you with serialization and schema evolution as described [here](#).

How can I access Producer and Consumer JMX metrics from Connect workers?

Add the following properties to the Connect worker config file. The `ConfluentMetricsReporter` metrics reporter will pull out producer and consumer JMX metrics from each worker JVM and write them into a Apache Kafka® topic (`_confluent-connect-metrics`).

```
metric.reporters=io.confluent.metrics.reporter.ConfluentMetricsReporter
confluent.metrics.reporter.bootstrap.servers=metrics-kafka:9092
confluent.metrics.reporter.topic=_confluent-connect-metrics
confluent.metrics.reporter.topic.replicas=1
confluent.metrics.reporter.whitelist=.*
confluent.metrics.reporter.publish.ms=60000
```

The messages in `_confluent-connect-metrics` can be consumed using the following command.

```
kafka-console-consumer --topic _confluent-connect-metrics \
  --bootstrap-server metrics-kafka:9092 \
  --formatter io.confluent.metrics.reporter.ConfluentMetricsFormatter
```

How can I use plain JSON data with Connect?

When using plain JSON data with Connect, users will see the following kind of error message:

```
org.apache.kafka.connect.errors.DataException: JsonSerializer with schemas.enable requires "schema" and "payload" fields and may not contain additional fields
```

You will need to set the `schemas.enable` parameters for the converter to `false` as described [here](#).

Does source connector X support output format Y?

Source connectors read data from a system and use a converter to change data from the connect data format being used to byte array. For source connectors, this means that any format for which a converter exists can be used to write data to Kafka. More details about converters are [here](#).

Why is CPU usage high for my Connect worker when no connectors have been deployed?

This is because upon startup the worker will read all files on the `CLASSPATH` to search for available connector plugins. Avoid having large portions of the file system on the worker's `CLASSPATH` to avoid this problem.

Can connect sink connectors read data written by other clients, e.g. a custom client?

This depends on how the data was written. The data must be written so that it is compatible with the converter specified in `key.converter` and `value.converter`. For example, applications writing Avro data using the converter provided in Schema Registry will produce compatible data to be read by a sink connector using the same converter.

After testing a connector in standalone mode, restarting it doesn't write the data again?

Standalone connectors store their offsets in a local file specified by `offset.storage.file.filename`. Connectors are designed to not reprocess data typically, so you would want to remove the file specified by this configuration to force the connector to reprocess the data.

Can I use a newer version of Connect with older brokers?

Connect follows the same [compatibility rules](#) as the Java clients for Kafka. As of Confluent Platform 3.2.0 and Kafka 0.10.2, in general it is possible to mix older and newer versions of both Kafka brokers and Kafka Connect workers. However, newer features in Kafka Connect (such as support for headers) will not work unless Connect is operating with a broker that also supports those features.

© Copyright 2019, Confluent, Inc. [Privacy Policy](#) | [Terms & Conditions](#). Apache, Apache Kafka, Kafka and the Kafka logo are trademarks of the [Apache Software Foundation](#). All other trademarks, servicemarks, and copyrights are the property of their respective owners.

[Please report any inaccuracies on this page or suggest an edit.](#)

1 Vote



Last updated on Oct 17, 2019.