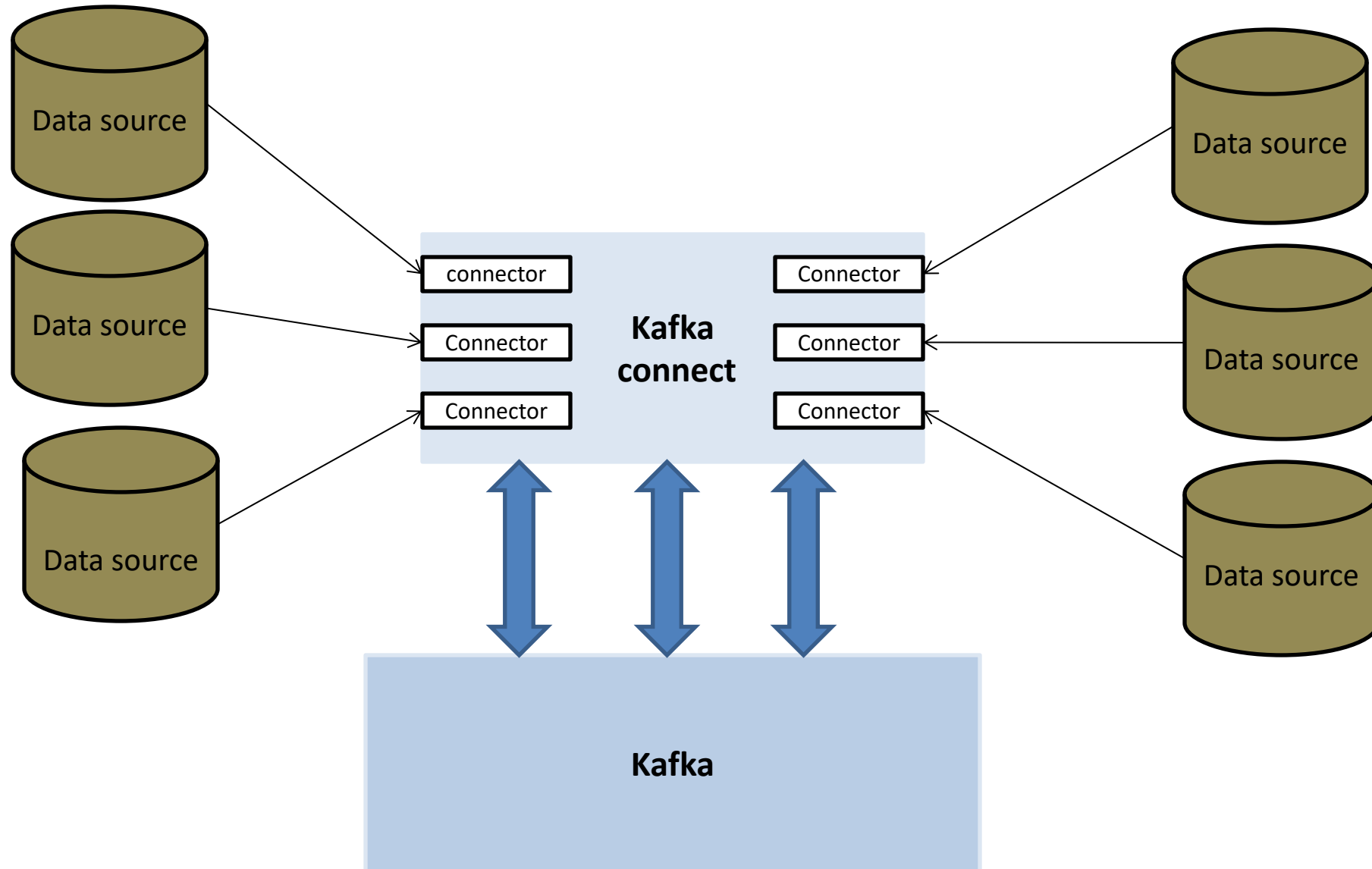


Kafka Connect

Simplified, scalable data import/export for Kafka

Kafka Connect : Source & Sink Connectors

Tos



Connectors and Tasks



JBDC SOURCE

S3 SINK

WORKER

S3 TASK #1

JBDC TASK #1

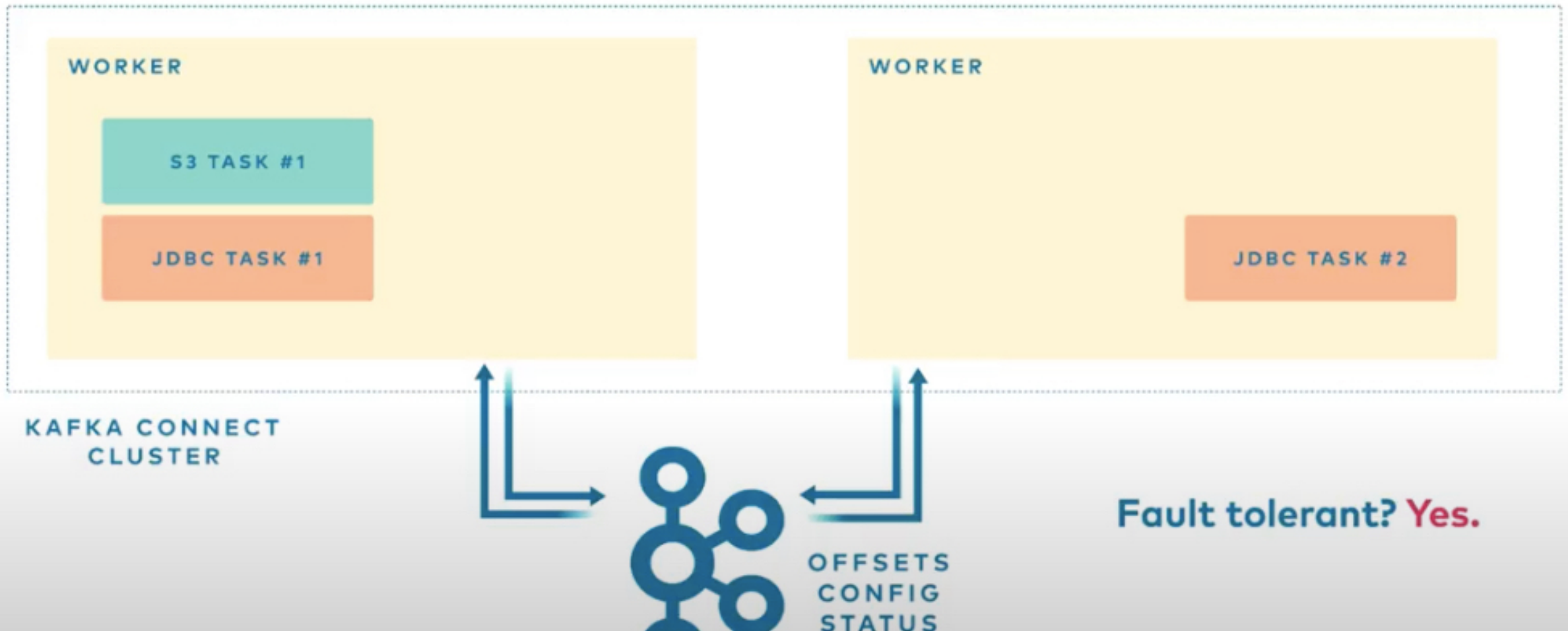
JBDC TASK #2

Kafka Connect Standalone Worker



Fault-tolerant? **No.**

Scaling the Distributed Worker



These parameters define the worker connection to the Kafka cluster.
See docs.confluent.io for a comprehensive list.

Parameter	Description
bootstrap.servers	A list of host/port pairs to use for establishing the initial connection to the Kafka Cluster.
Key.converter	Converter class for key connect data.
value.converter	Converter class for value connect data

- Starting Connect in standalone mode involves starting a process with one or more connect configurations:

❖ *Connect-standalone worker.properties connector1.properties [connector2.properties connector3.properties ...]*

- Each connector instance will be run in its own thread.
- Configuration will be covered later.

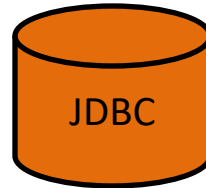
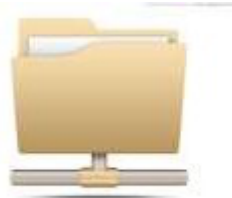
- Starting Connect in distributed mode involves starting connect on each worker node with the following:

connect-distributed worker.properties

- Connectors are then added, modified, and deleted via a REST API.
Configuration will be covered later.

Method	Path	Description
GET	/connectors	Get a list of active connectors.
POST	/connectors	Create a new connector
PUT	/connectors/(string:name)/config	Create a new connector, or update the configuration of an existing connector
GET	/connectors/(string: name)/config	Get configuration info for a connector.
GET	/connectors/(string: name)/tasks/<tasks-id>/status	Retrieve details for specific tasks
DELETE	/connectors/(string:name)	Delete a configured connector from the worker pool

- ❖ Confluent-supported connectors (included in Confluent Platform)



- ❖ Community-written connectors (just a sampling)



Lab : Kafka Connector – File connector