Transferring Data with Kafka Connect

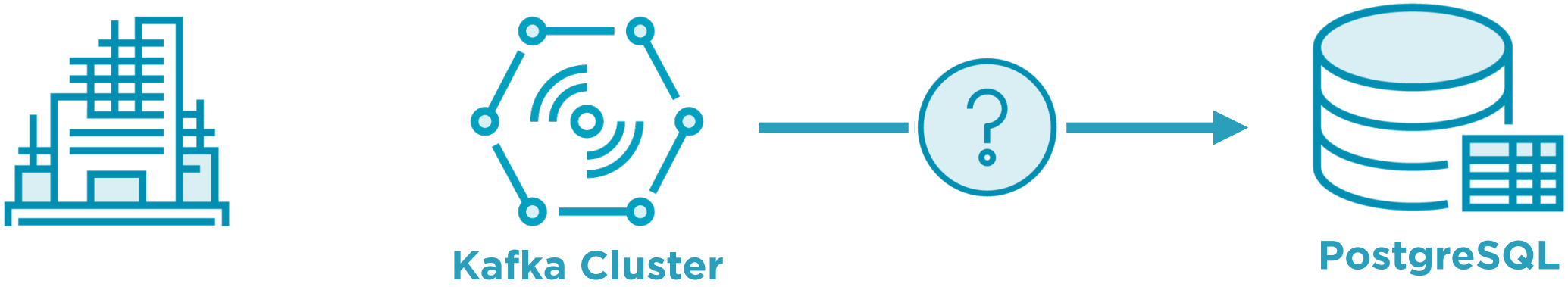
Bogdan Sucaciu
SOFTWARE ENGINEER @ AXUAL

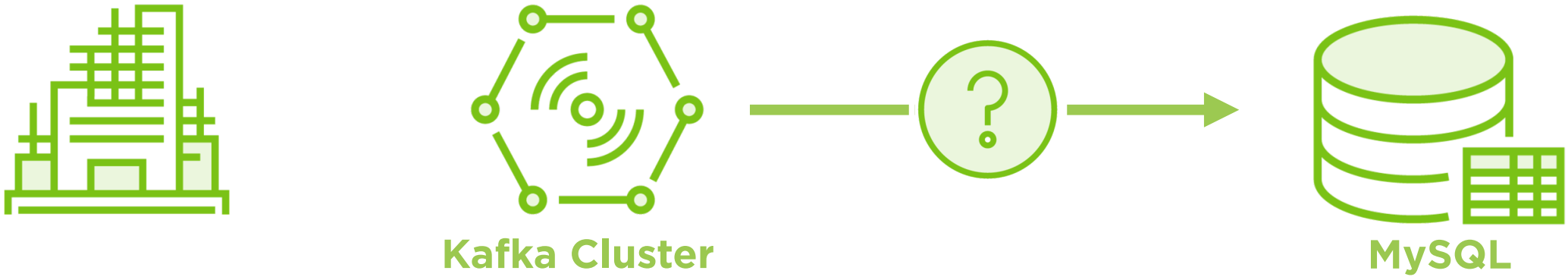
@BSucaciu

Why Kafka Connect?

Code Reuse

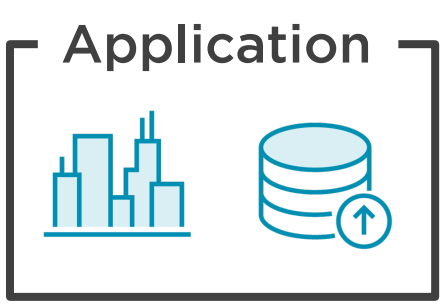
A Tale of Two Companies





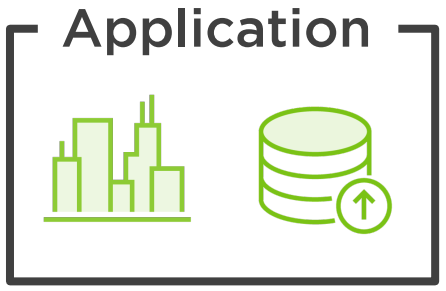
A Tale of Two Companies

Kafka Cluster



PostgreSQL

Kafka Cluster



MySQL

Limited Number of Technologies

MongoDB

Neo4J

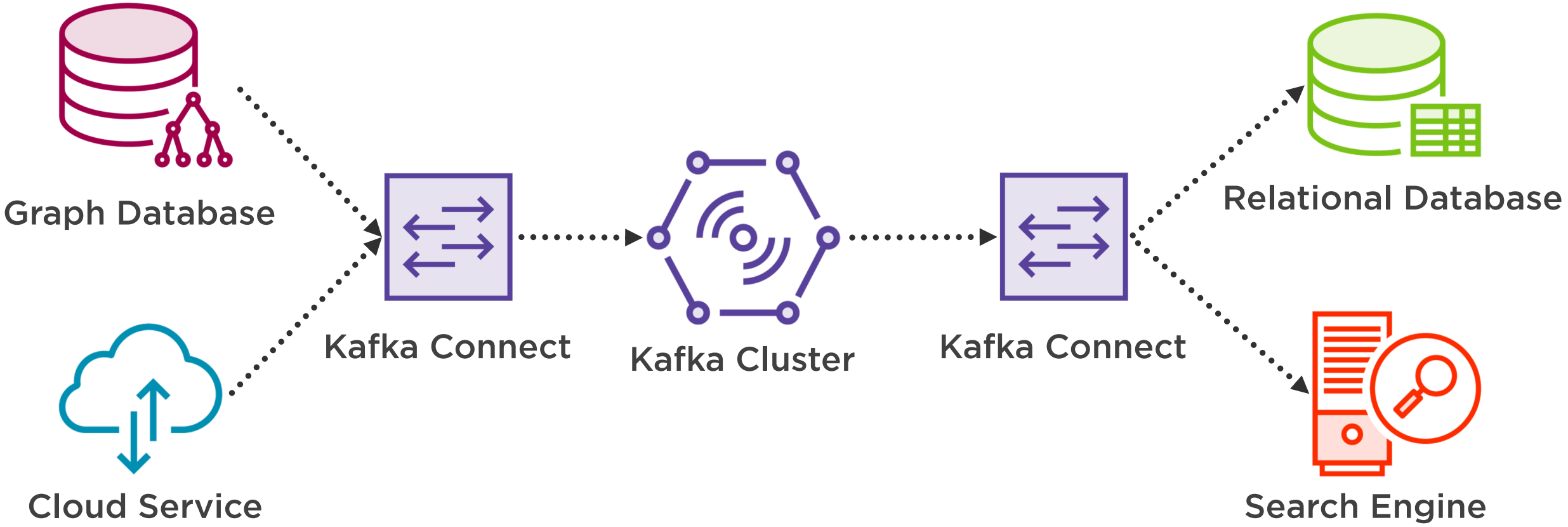
ElasticSearch

Solr

AWS S3

Twitter

Kafka Connect

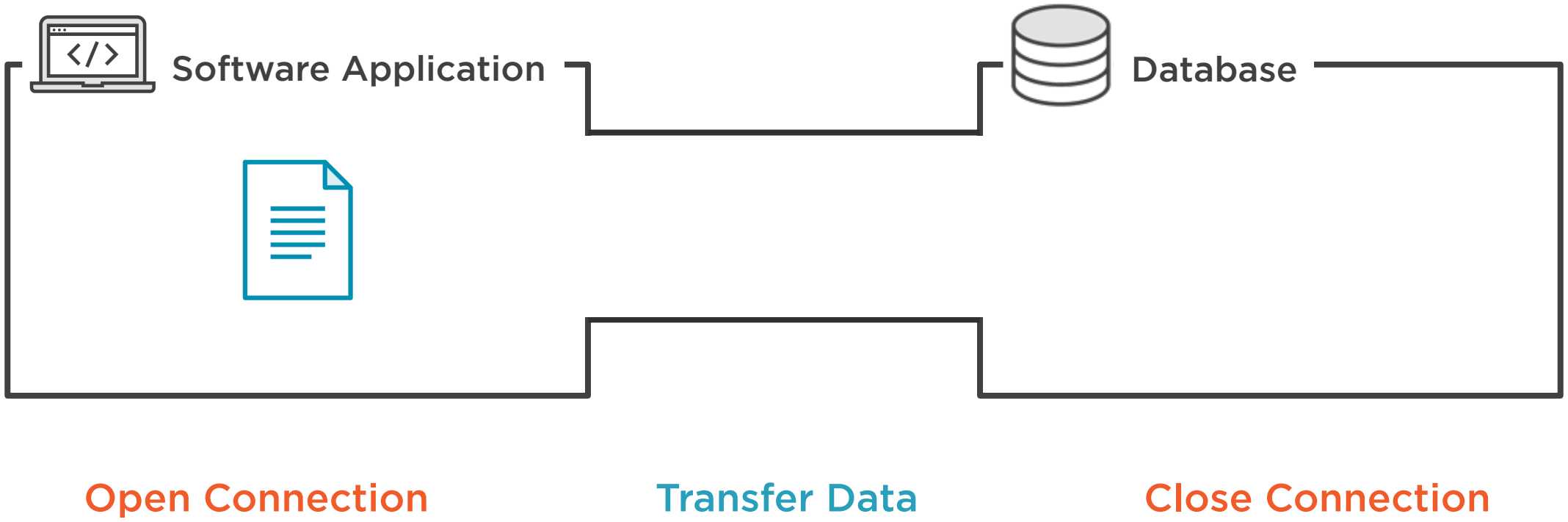


Why Kafka Connect?

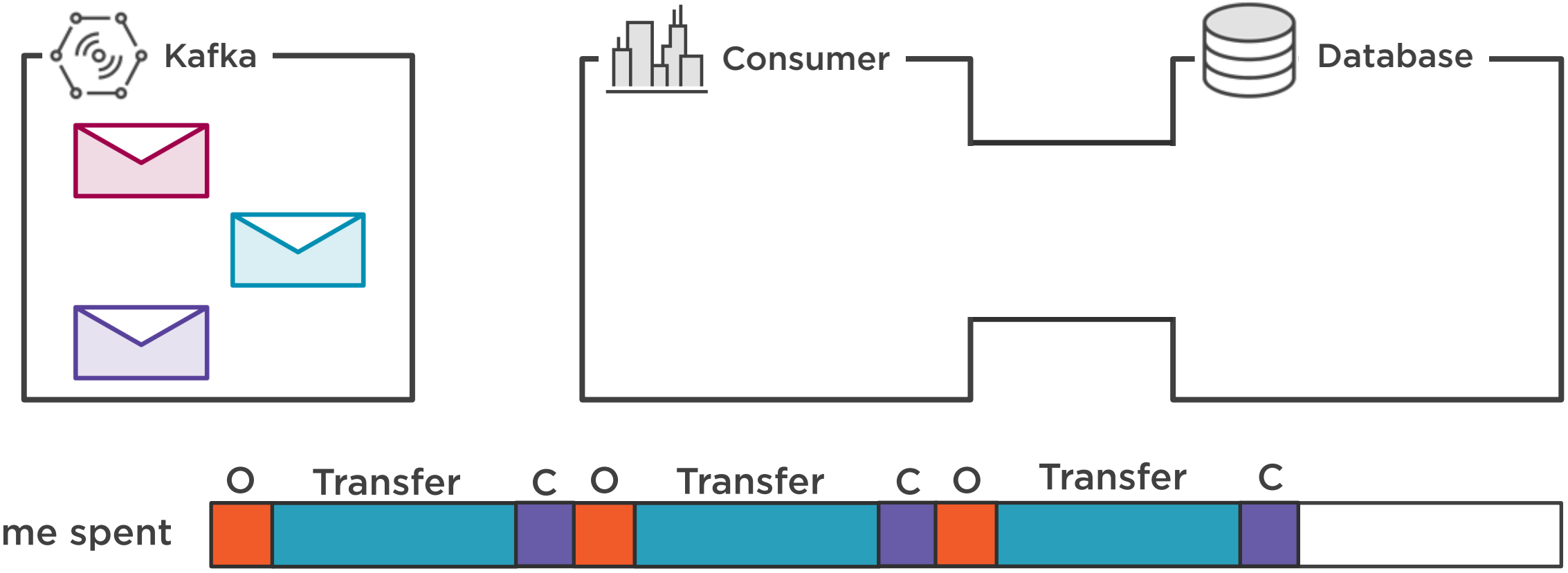
Code Reuse

Avoid reinventing the wheel

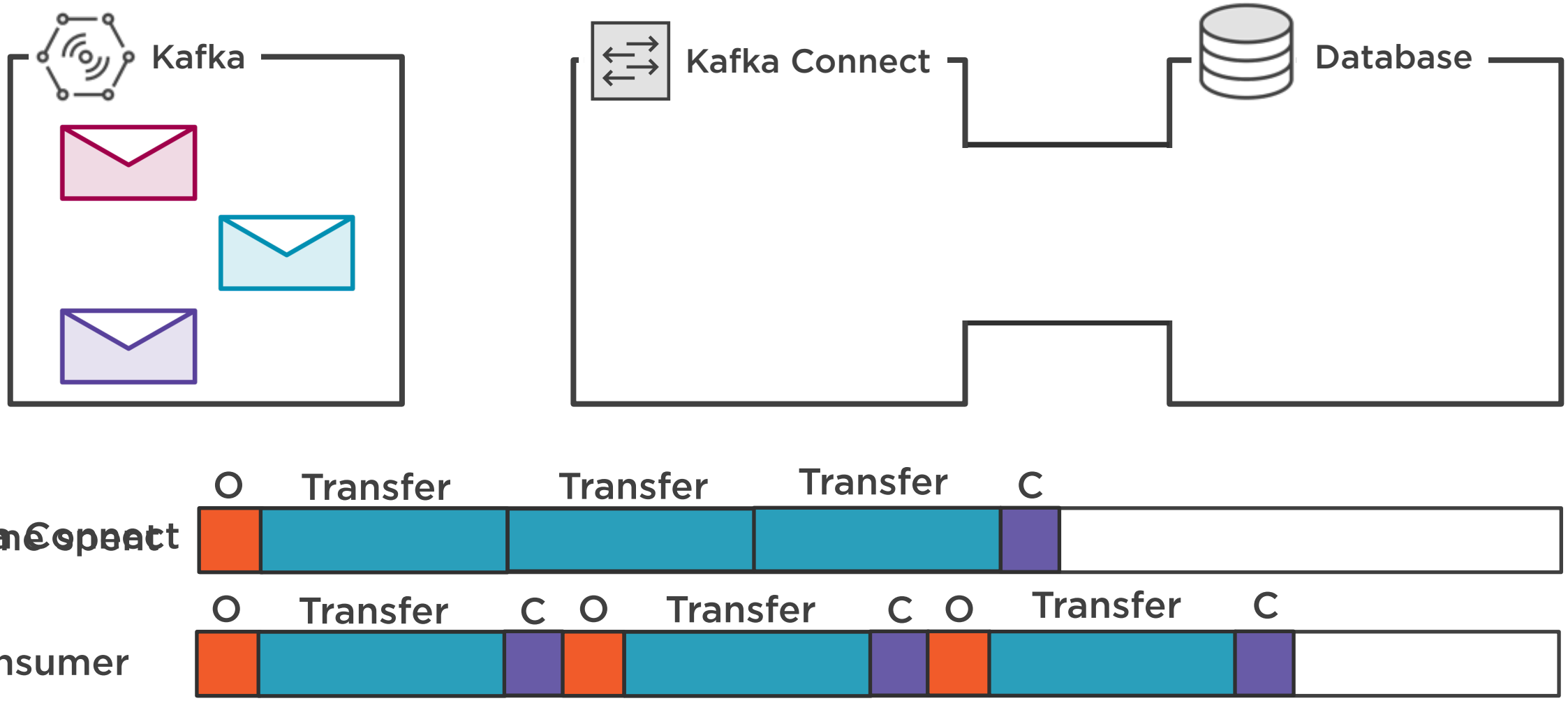
Connecting to a Database



Transferring Data to a Database



Transferring Data to a Database



Why Kafka Connect?

Code Reuse

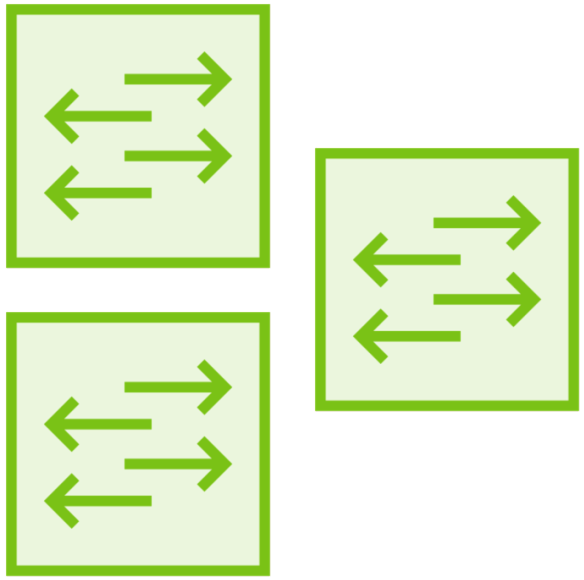
Avoid reinventing the wheel

Scalability

Scalability

Standalone

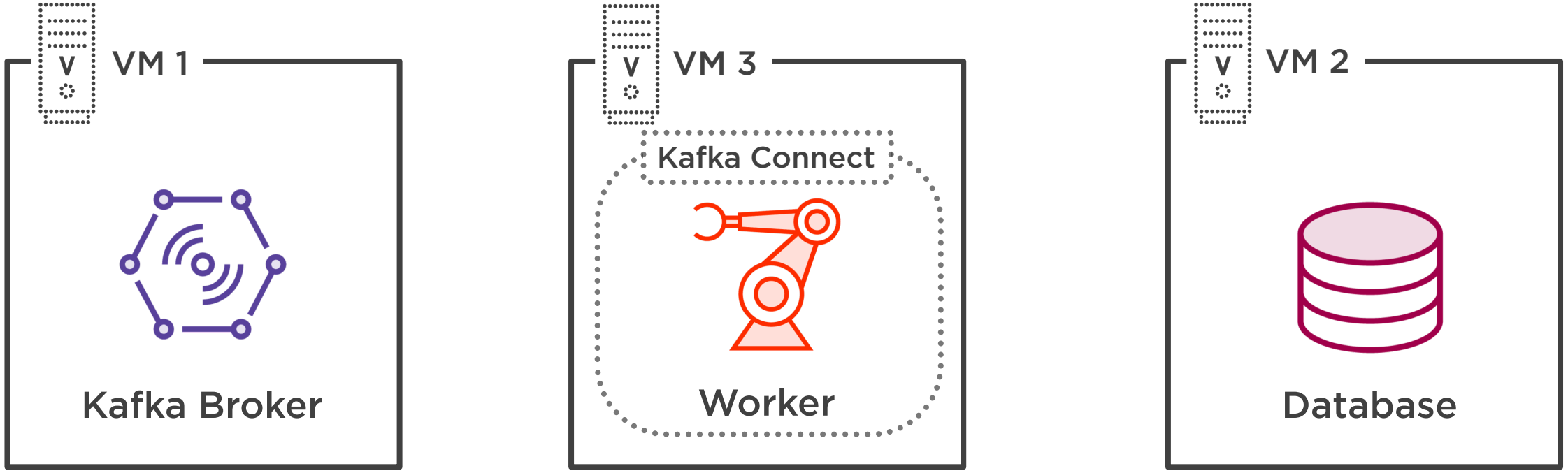
A single instance of Kafka Connect



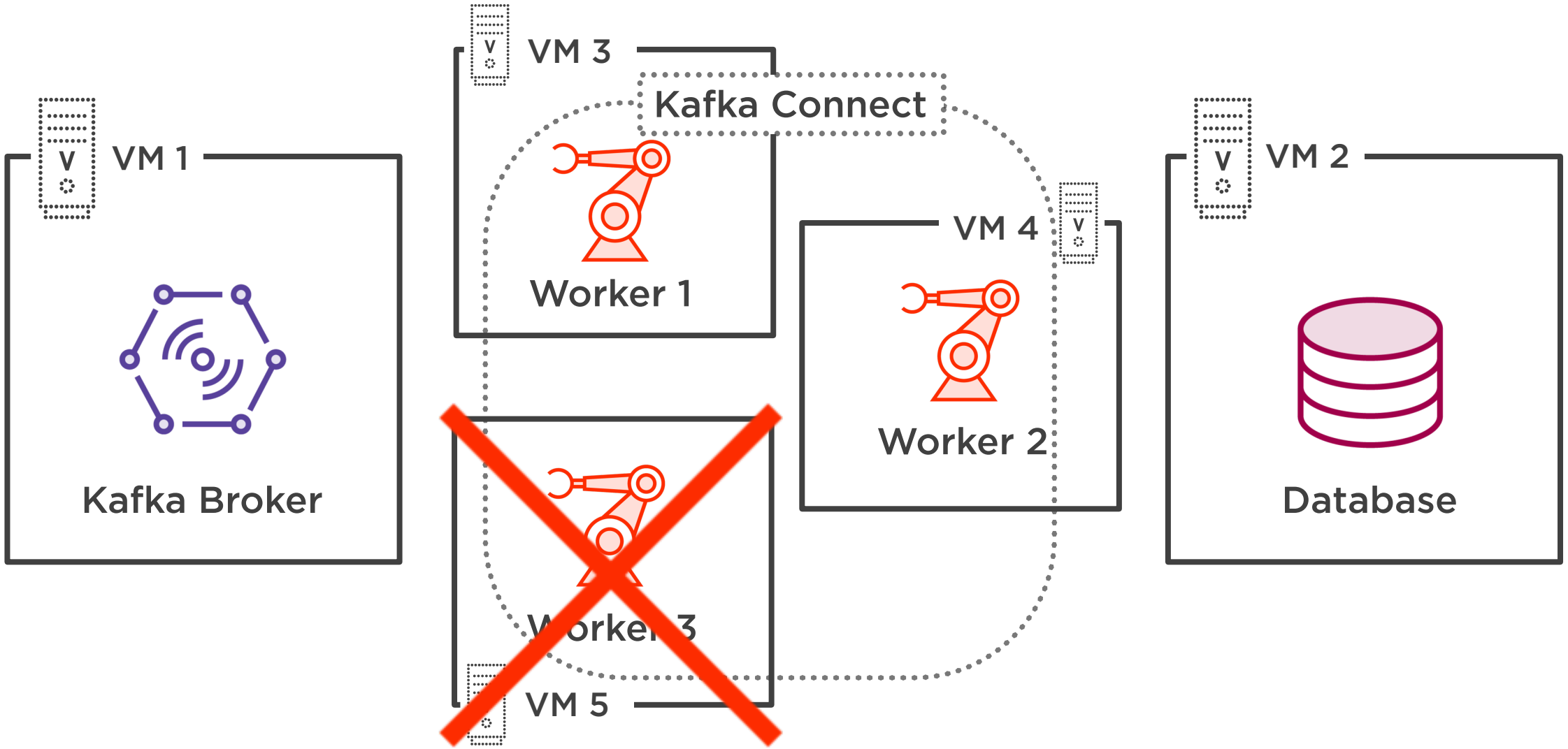
Distributed

Multiple instances on separate machines

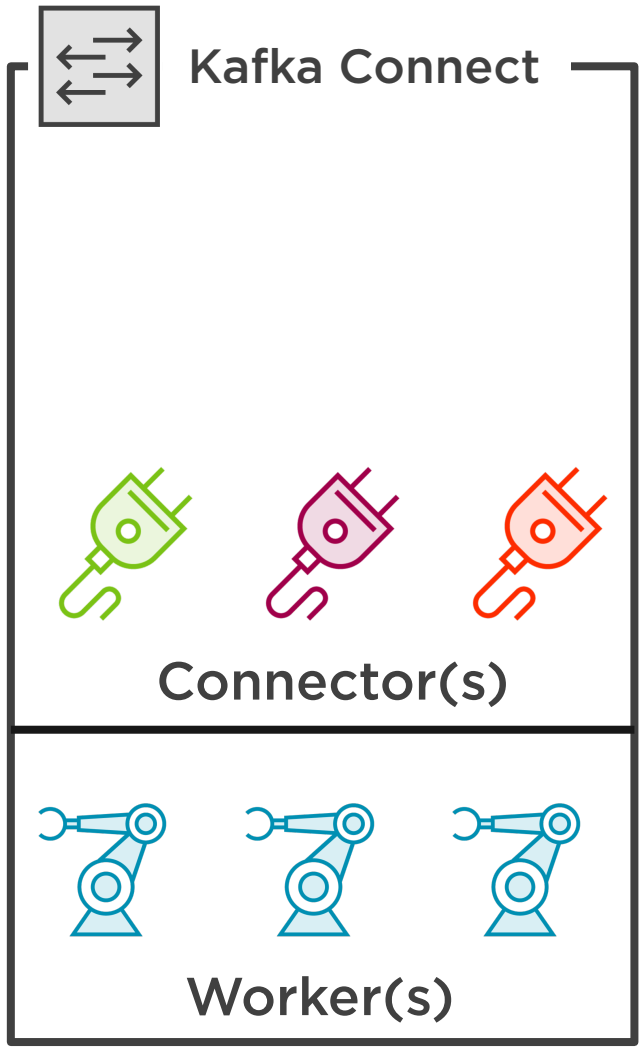
Kafka Connect Architecture



Kafka Connect Architecture



Kafka Connect Architecture

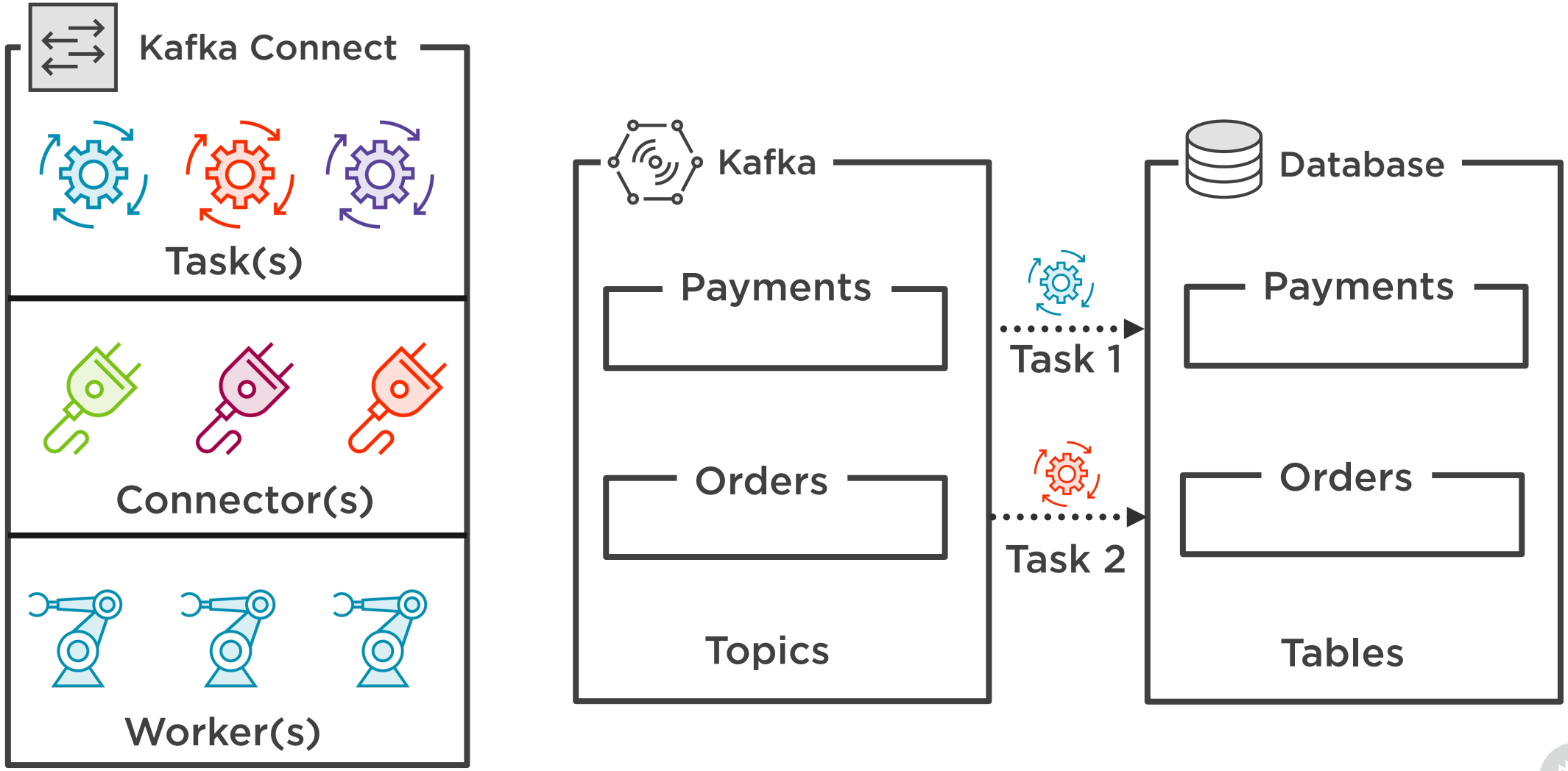


Relational Database

Graph Database

Search Engine

Kafka Connect Architecture



Connectors

Source Connectors

Transfer data
from a Source to Kafka

Sink Connectors

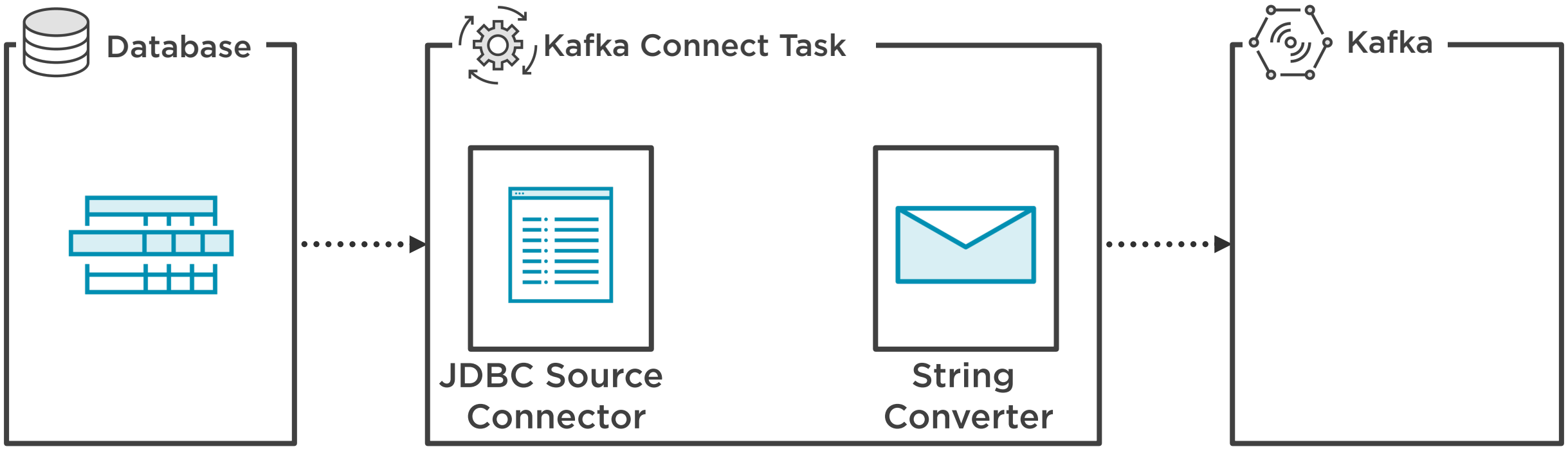
Transfer data
from Kafka to a Sink

Connectors

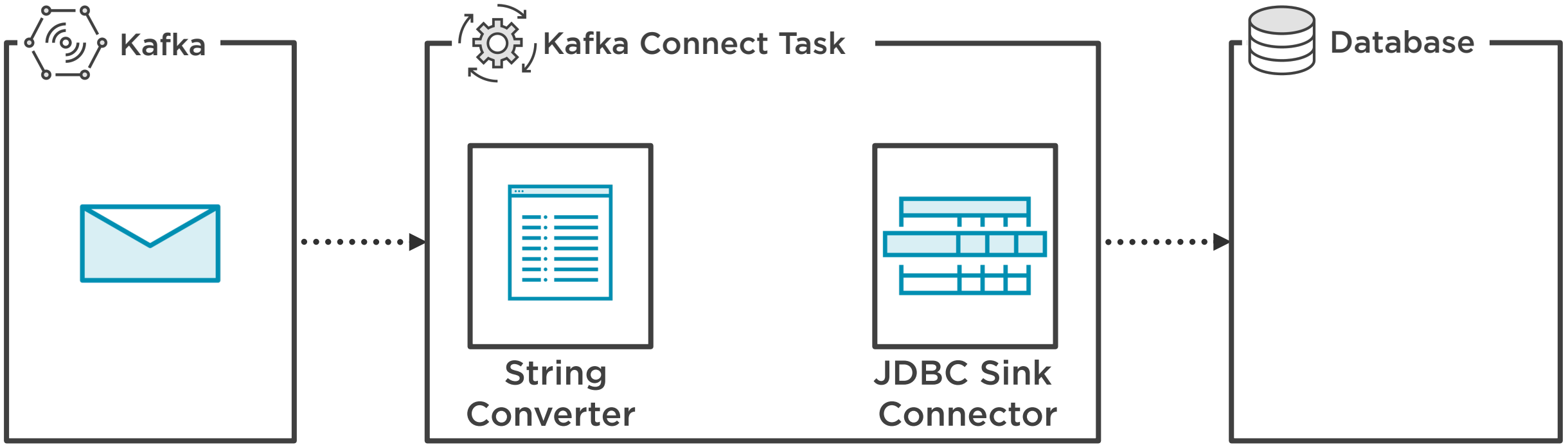

Database
JDBC Source Connector
Kafka Broker


Kafka Broker
AWS S3 Sink Connector
AWS S3

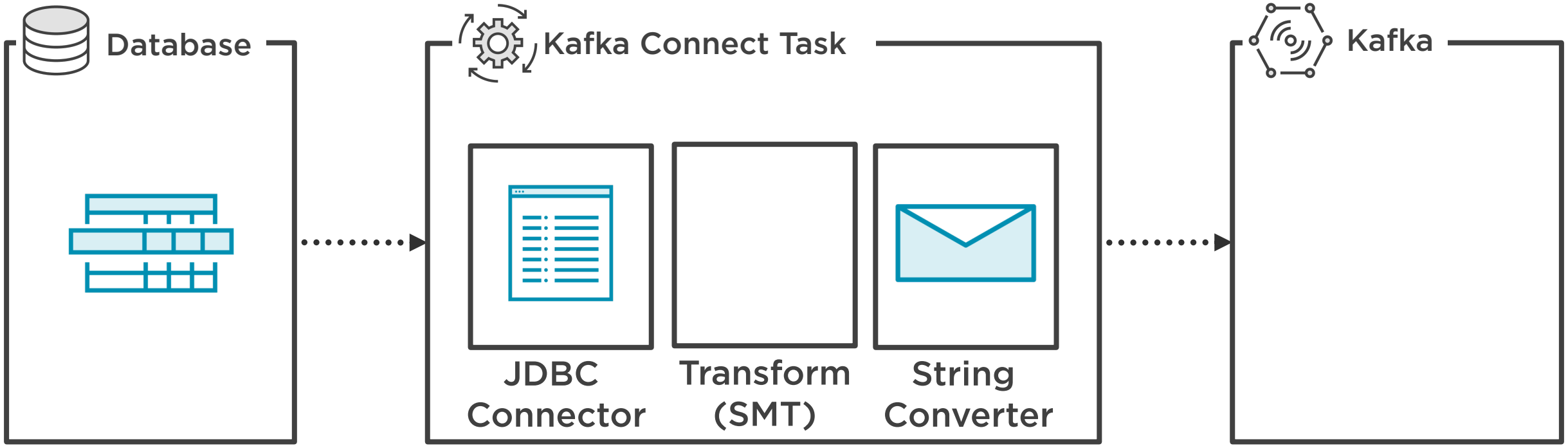
Converters


Database
Kafka Connect Task
Kafka
JDBC Source Connector
String Converter

Converters

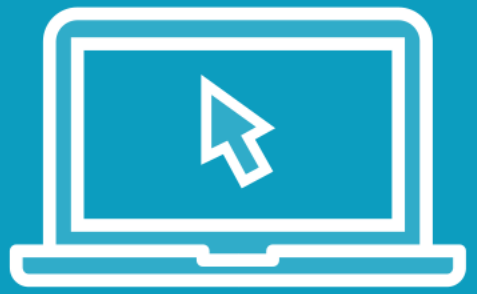

Kafka
Kafka Connect Task
Database
String
Converter
JDBC Sink
Connector

Transforms



SMT – Simple Message Transforms

Demo



Kafka Connect in action

JDBC Sink Connector

Prerequisites:

- MySQL DB
- bit.ly/mysql-jdbc-connector
- bit.ly/mysql-driver

Summary

Why Kafka Connect?

How it works

Capabilities