



Advanced
Consulting
Partner

Data & Analytics
Competency

Solution Provider

Public Sector Partner

Immersion Day Partner

Apache Kafka (Confluent) / Attunity Replicate

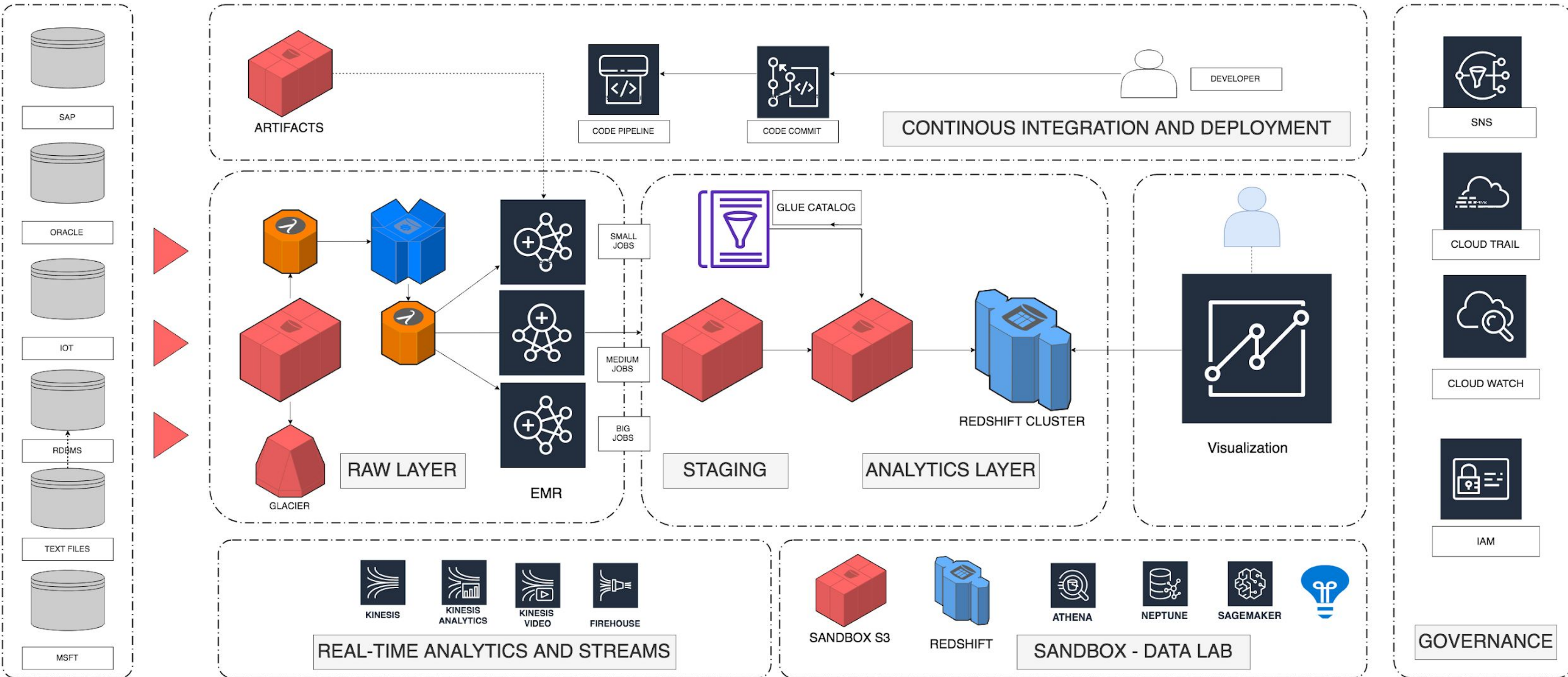
We Love the Cloud

www.morrisopazo.com

contacto@morrisopazo.com

- ¿Qué es Apache Kafka?
- Terminología Apache Kafka
- ¿Qué es Confluent?
- ¿Qué es Attunity Replicate?
- Attunity Replicate & Apache Kafka
- Instalación Confluent
- Configuración Confluent
- Creación Tópicos
- Instalacion Attunity
- Creación Endpoint
- Creación Tarea
- Configuración de Tarea

ARQUITECTURA REFERENCIAL DE SOLUCIÓN BIG DATA

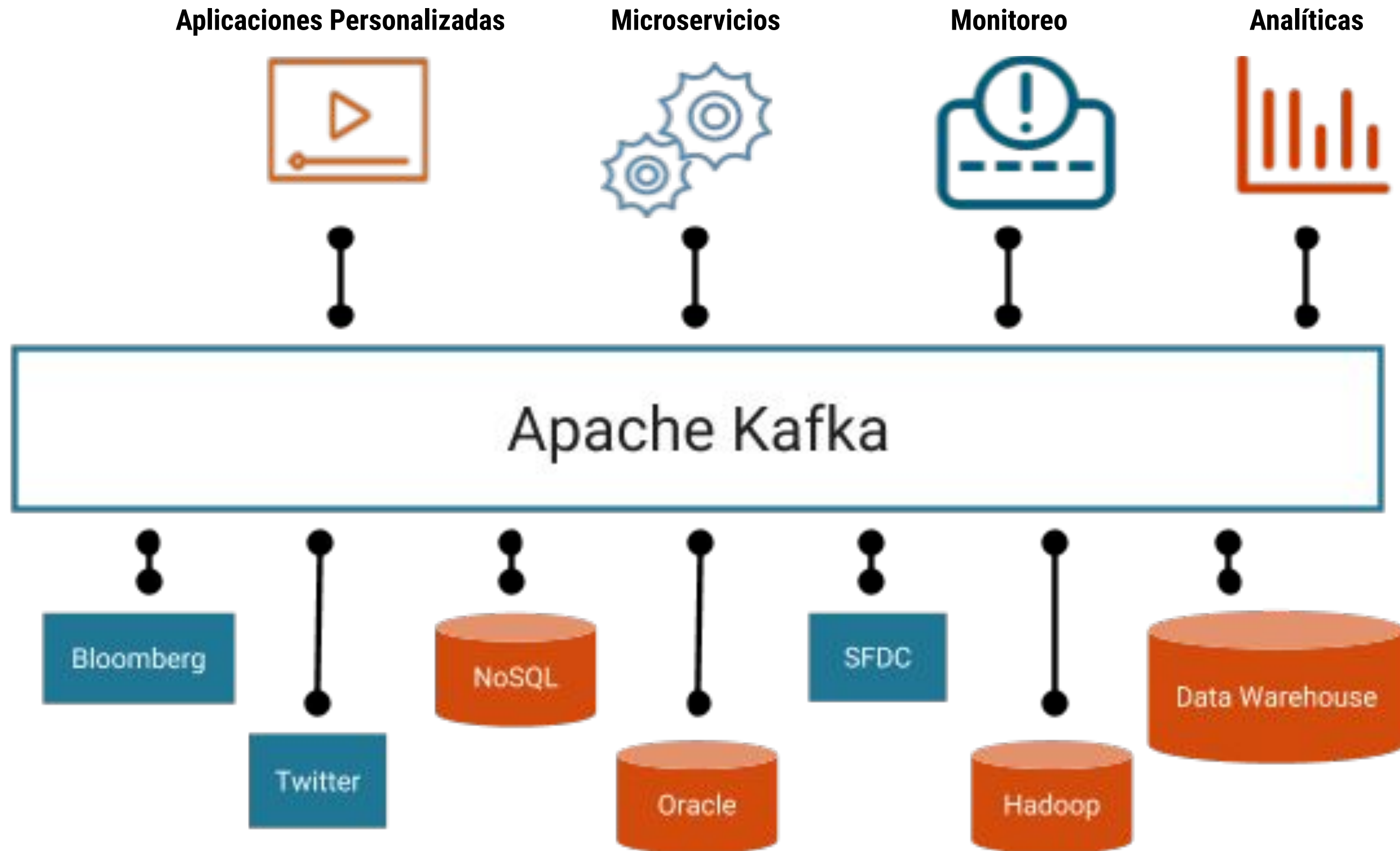


¿QUÉ ES APACHE KAFKA?



- Apache Kafka es una plataforma de transmisión distribuida de datos en tiempo real.
- Permite almacenar flujos de registros de una manera duradera y tolerante a fallas.
- Procesa flujos de registros a medida que ocurren.
- Alto rendimiento: diseñado para una gran cantidad de mensajes pequeños
- Se utiliza para construir pipelines de datos de transmisión en tiempo real y aplicaciones de transmisión

¿QUÉ ES APACHE KAFKA?



TERMINOLOGÍA APACHE KAFKA



Broker	Servidor Kafka. Varios brokers forman un cluster.
Cluster	Grupo de dos o más brokers que brindan redundancia y escalabilidad
Productor	Cliente Kafka. Es quien envía los mensajes al broker.
Consumidor	Cliente Kafka. Es quien recibe los mensajes desde el broker.
Tópico	Un topic es una categoría en la que se publican los registros. Cada Topic de Kafka puede tener desde cero a muchos consumidores que se suscriben a los datos escritos en él.
Partición	Subconjunto de topics creado para permitir la redundancia y lectura / escritura paralela para un mayor rendimiento
Mensaje	Cualquier búfer (texto / binario). El formato debe coordinarse entre Productores y Consumidores
Clave del Mensaje	Opcional. Utilizado para particionamiento y compactación.
Offset	Posición de un mensaje en una partición. Utilizado por el consumidor.

- Confluent Control Center es un sistema basado en GUI para administrar y monitorear Kafka.
- Esta plataforma permite administrar fácilmente Kafka Connect, crear, editar y administrar conexiones a otros sistemas.
- También permite monitorear los flujos de datos del productor al consumidor, asegurando que se entreguen todos los mensajes y midiendo cuánto tiempo lleva entregarlos.

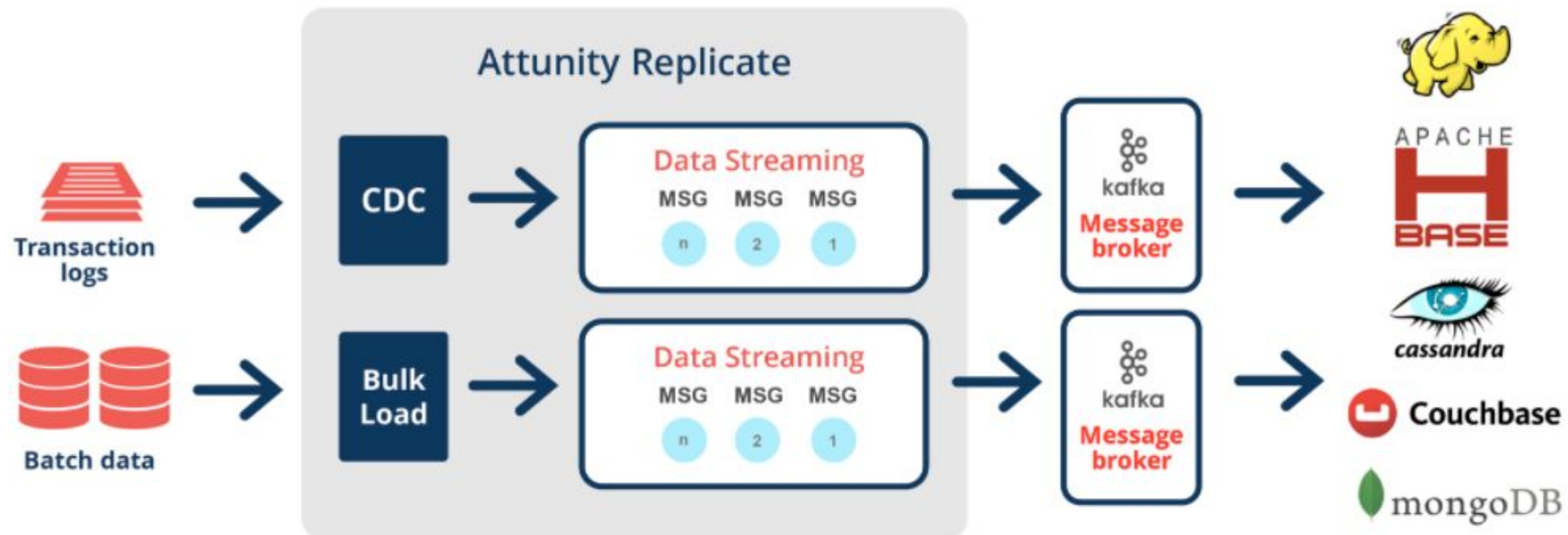
¿QUÉ ES ATTUNITY REPLICATE?

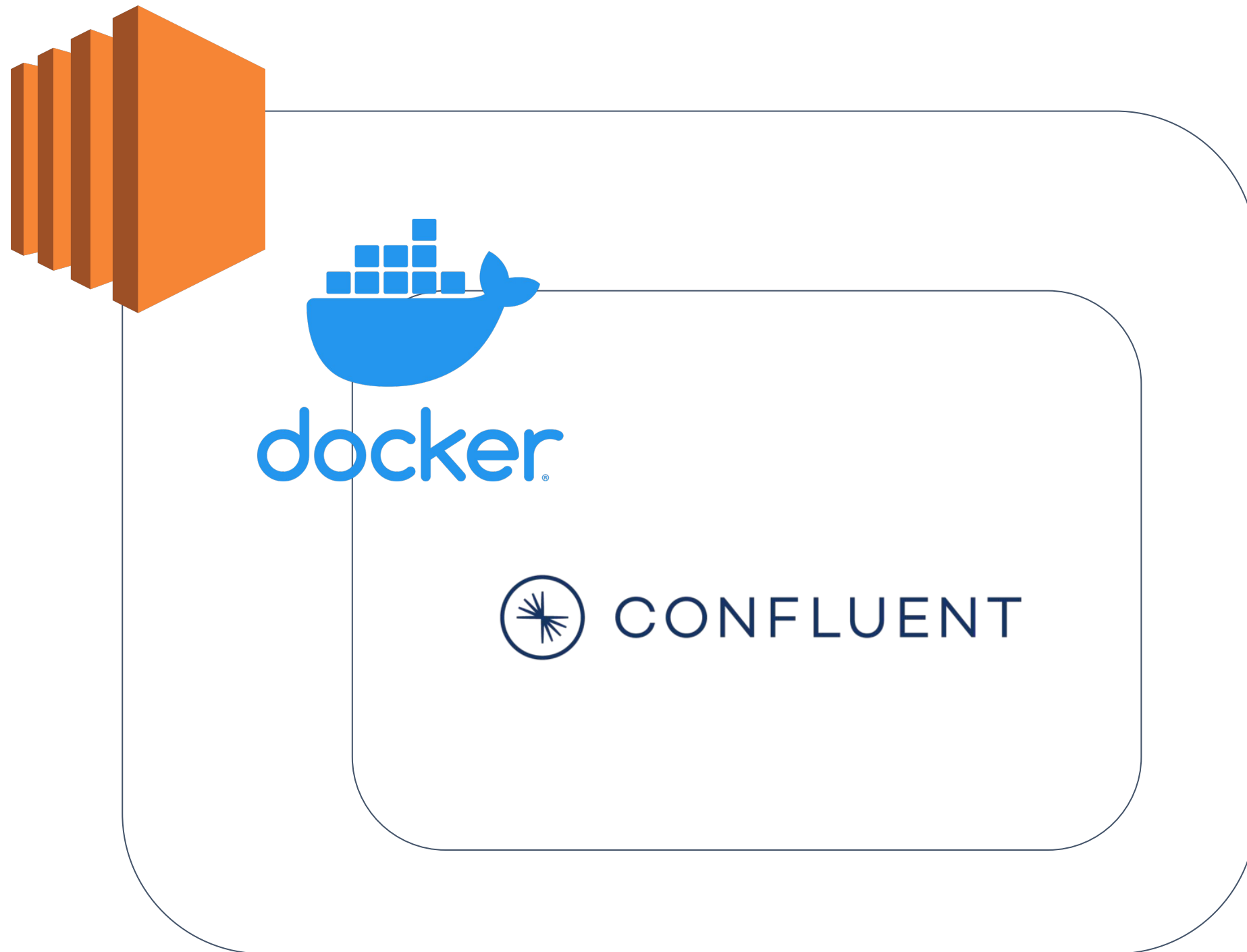


- Es una herramienta que permite a las empresas acelerar la replicación, ingesta y transmisión de datos en una gran variedad de bases de datos, data warehouses y plataformas de Big Data.
- Facilita el mover grandes volúmenes de datos con seguridad, eficacia, y con un impacto muy bajo en las operaciones.
- Simplifica la ingesta masiva en plataformas de Big Data desde miles de fuentes.

ATTUNITY REPLICATE & APACHE KAFKA

- Attunity Replicate actúa como productor de Apache Kafka
- Attunity Replicate permite la conexión con los brokers de Apache Kafka y enviar mensajes a un tópico específico.
- Permite la configuración de endpoints para Apache Kafka





- EC2 Amazon Linux 2 AMI
t3.xlarge (16GB RAM)
- Requisitos mínimos 8GB RAM
asignados a Docker
- <http://<HOST>:9021>

- Se debe configurar el cluster para poder recibir información desde el exterior. para esto se debe modificar la propiedad **advertised.listeners** en los ajustes del cluster.

```
PLAINTEXT://broker:29092,PLAINTEXT_HOST://[IP o DNS]:9092
```

CREACIÓN TÓPICOS

 CONFLUENT


Cluster 1

CONTROLCENTER.CLUSTER

Overview

Brokers

Topics

Connect

ksqlDB

Consumers

Replicators

Cluster settings

All topics

☒ Hide internal topics

+ Add a topic

Topics	Availability			Throughput	
Topic name	Under replicated partitions	Out of sync followers	Out of sync observers	Bytes/sec produced	Byt
docker-connect-configs	0 of 1	0 of 1	0 of 0	0	0
docker-connect-offsets	0 of 25	0 of 25	0 of 0	--	0
docker-connect-status	0 of 5	0 of 5	0 of 0	--	0
world_data	0 of 1	0 of 1	0 of 0	0	0

ALL TOPICS >

New topic

Topic name*

nombre topico

Number of partitions* ⓘ

1 ▼

Create with defaults

Customize settings

[Cancel](#)

Otras Configuraciones:

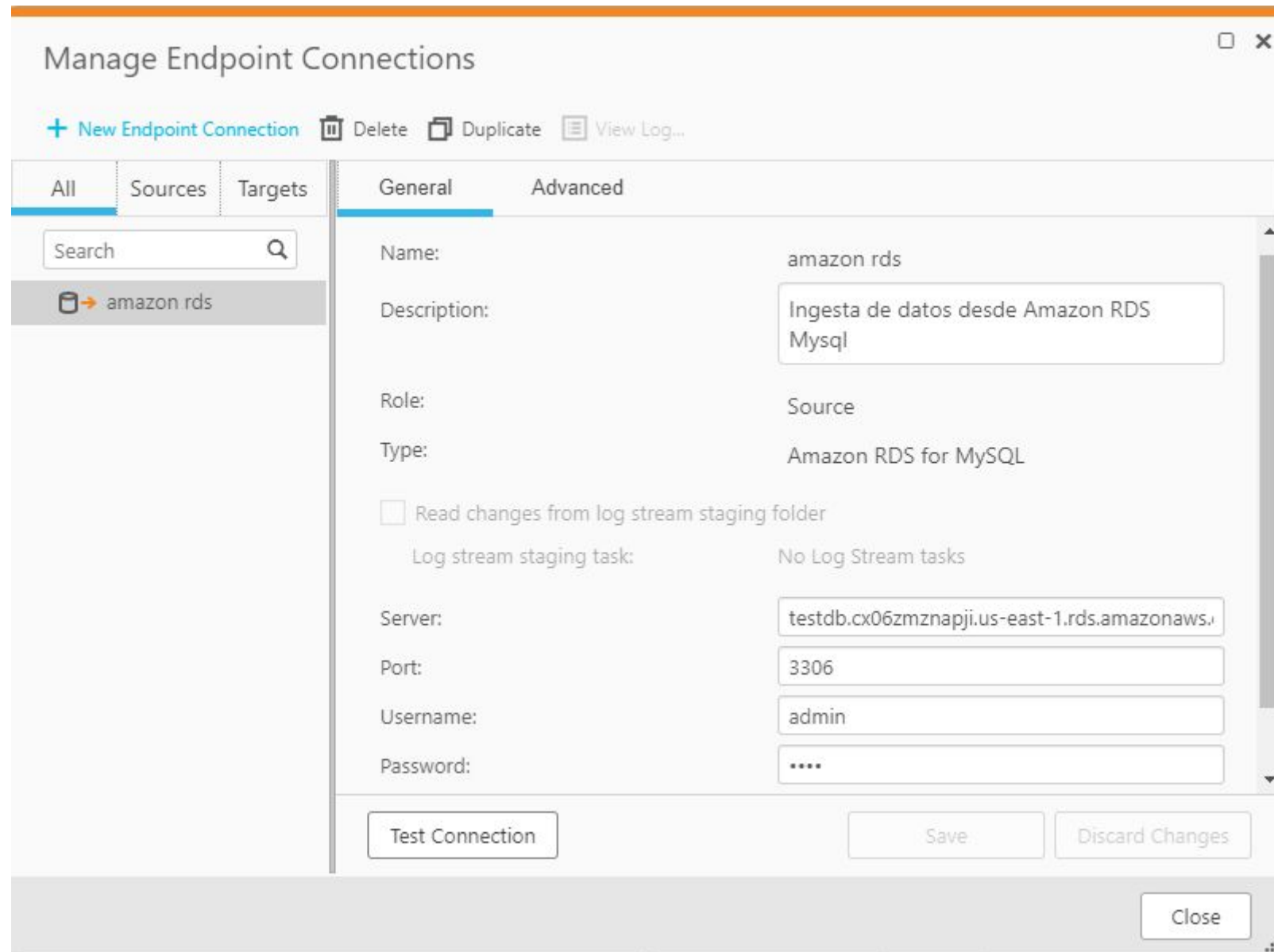
- Disponibilidad
- Política de limpieza
- Tamaño máximo en el disco
- Tiempo de retención

- Free Trial disponible en página web
- Sesiones 4 horas
- No requiere configuración



- Attunity Replicate permite acceder o enviar datos mediante conectores que pueden ser de tipo **source** o tipo **target**.
- Para manejar estos conectores seleccionar la opción **Manage Endpoint Connections**
- Attunity soporta gran cantidad de tipos de conexiones

CREACIÓN ENDPOINT

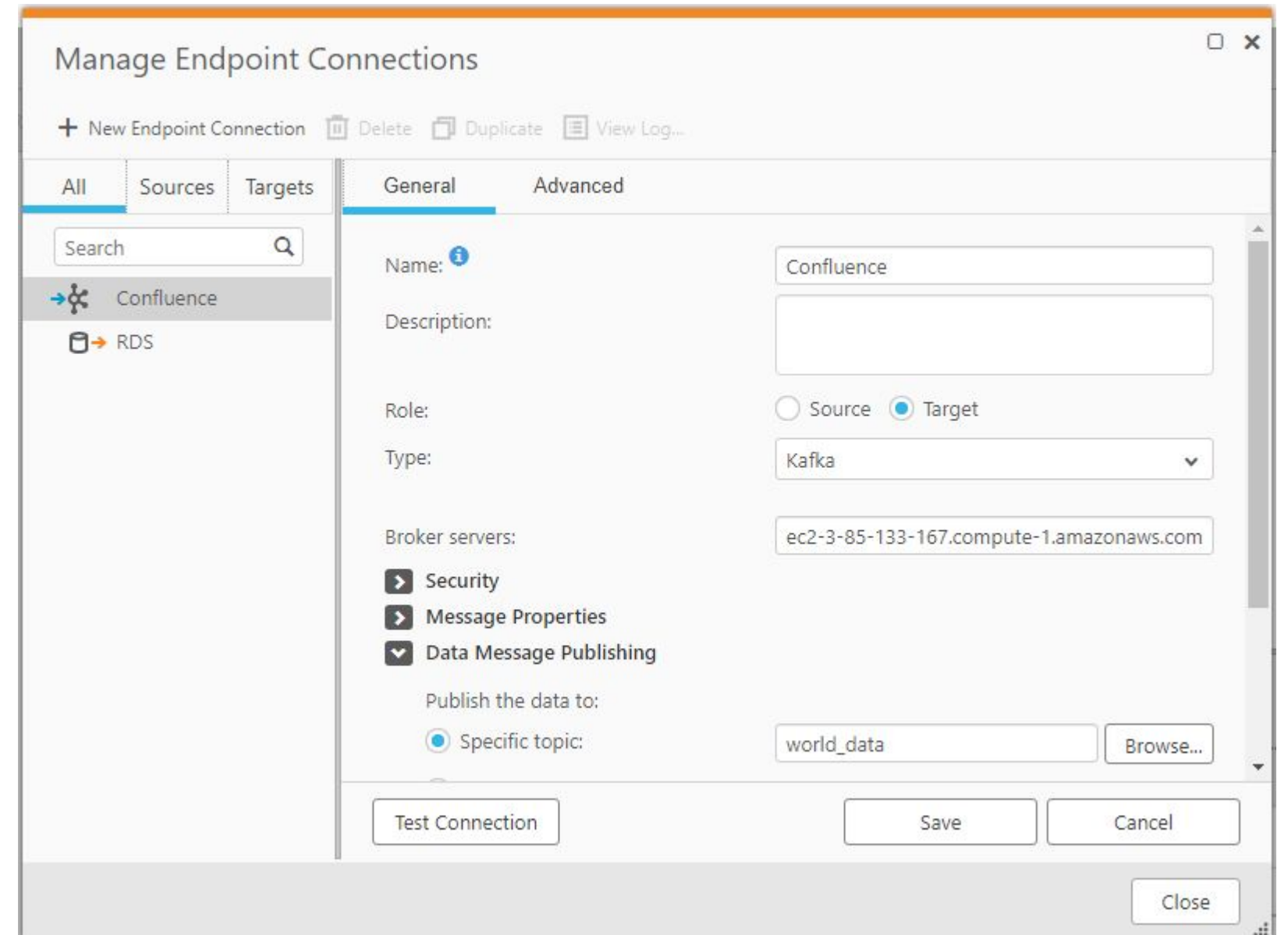


The screenshot shows the 'Manage Endpoint Connections' dialog with the 'General' tab selected. The left sidebar lists 'amazon rds' as the selected source. The main form fields are as follows:

Field	Value
Name	amazon rds
Description	Ingesta de datos desde Amazon RDS Mysql
Role	Source
Type	Amazon RDS for MySQL
Read changes from log stream staging folder	☐
Log stream staging task	No Log Stream tasks
Server	testdb.cx06zmznapij.us-east-1.rds.amazonaws.com
Port	3306
Username	admin
Password	

Buttons at the bottom: Test Connection, Save, Discard Changes, Close.

Source Endpoint



The screenshot shows the 'Manage Endpoint Connections' dialog with the 'General' tab selected. The left sidebar lists 'Confluence' as the selected target. The main form fields are as follows:

Field	Value
Name	Confluence
Description	
Role	☒ Target
Type	Kafka
Broker servers	ec2-3-85-133-167.compute-1.amazonaws.com
Security	☒
Message Properties	☒
Data Message Publishing	☒
Publish the data to:	☒ Specific topic: world_data

Buttons at the bottom: Test Connection, Save, Cancel, Close.

Target Endpoint

- Permite replicar datos desde un punto de origen hacia un punto objetivo.
- Endpoint Target y Endpoint Source
- Para crear una nueva tarea se debe presionar la opción New Task desde la pantalla principal



New Task

Name

RDSToConfluent

Description

Permite enviar mensajes desde una base de datos RDS Mysql a Un cluster de kafka.

Replication Profile

Select the profile that best describes your replication use case

☒ Unidirectional

☐ Bidirectional

☐ Log Stream

Task Options

Select the replication options for this task



Full Load



Apply Changes



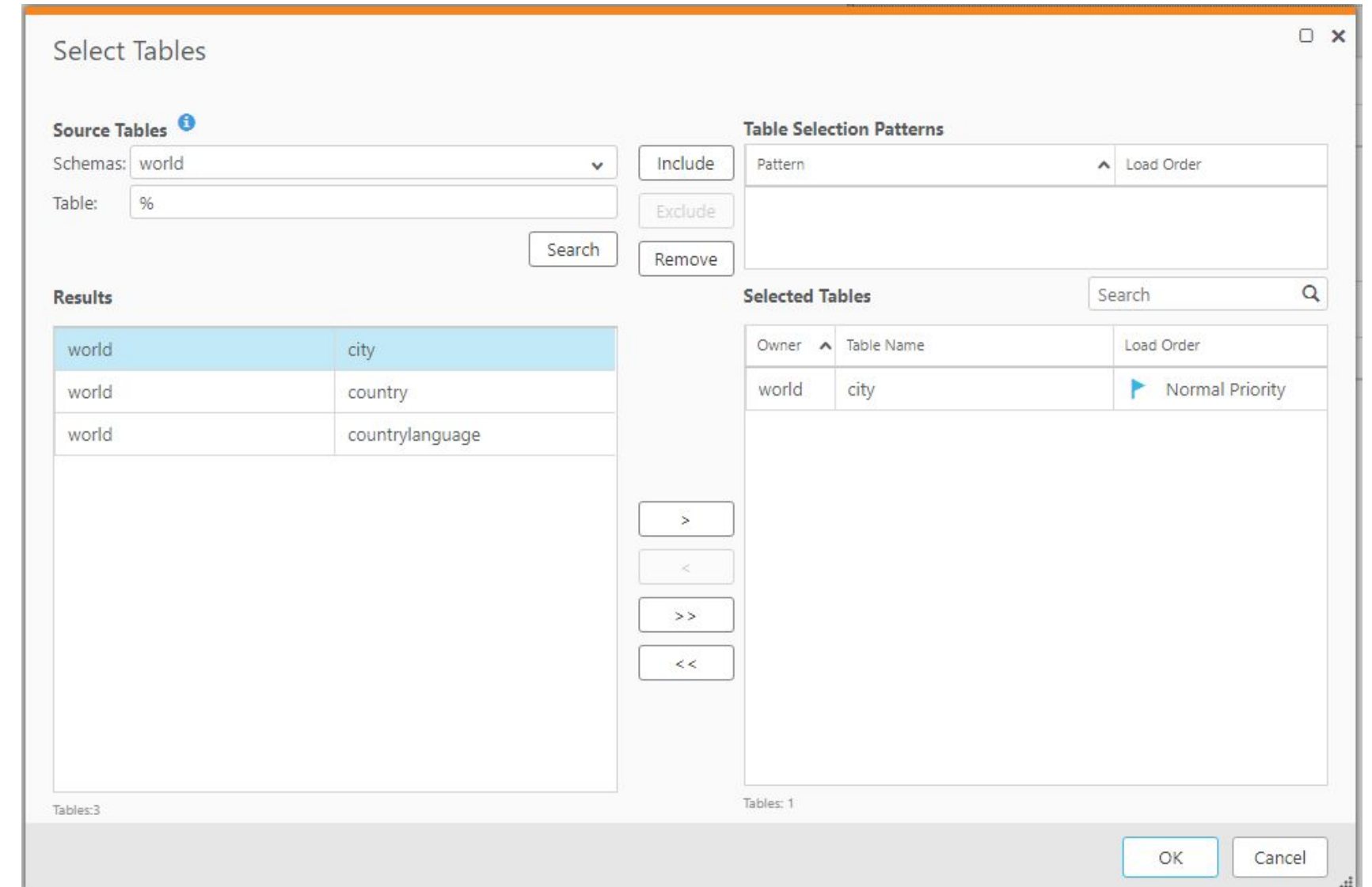
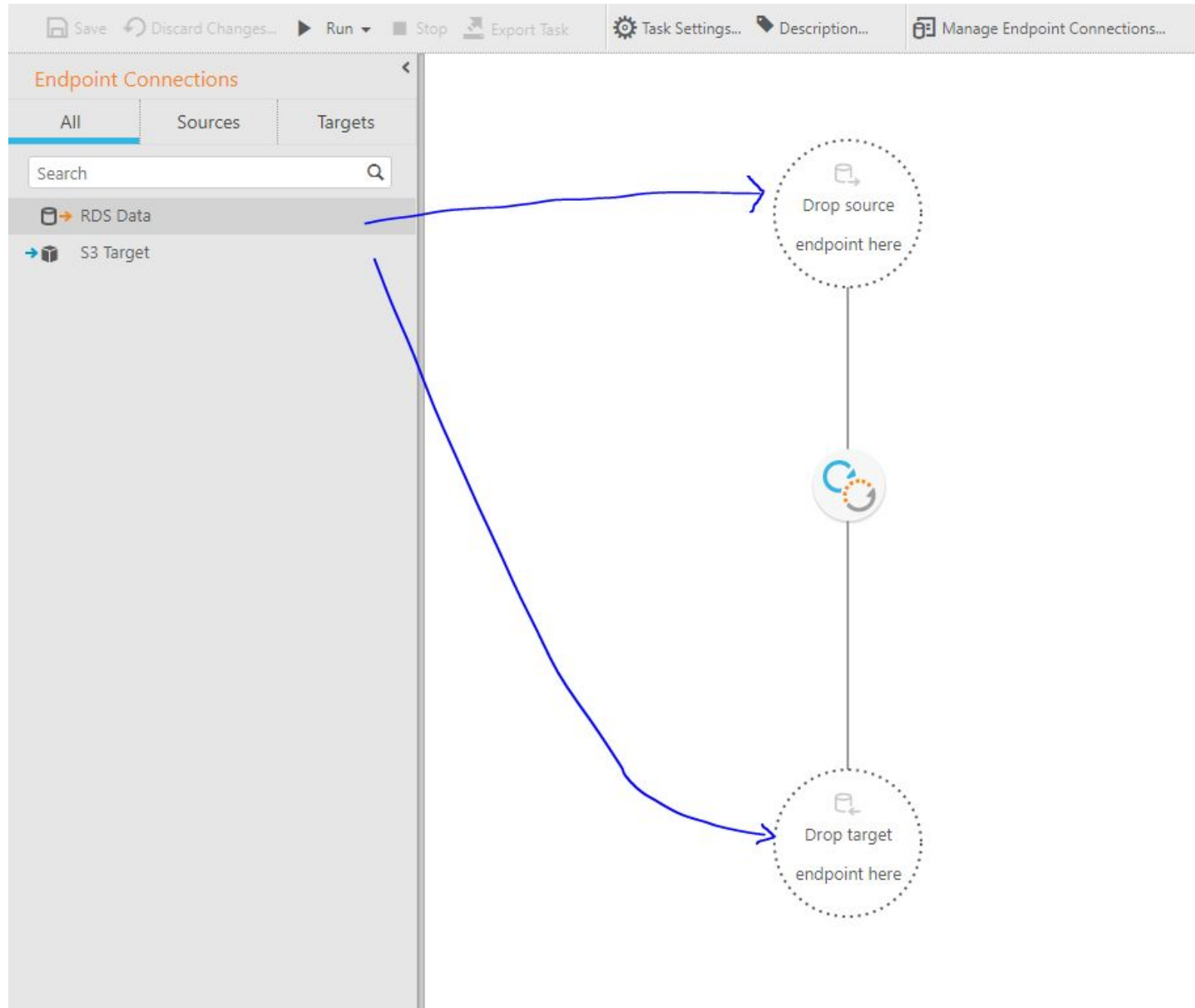
Store Changes

Perform a full load of target tables, then keep them up-to-date by applying changes.

OK

Cancel

CONFIGURACIÓN DE TAREA



Select Tables

Source Tables ⓘ

Schemas: world

Table: %

Include

Exclude

Search

Remove

Table Selection Patterns

Pattern Load Order

Results

Owner	Table Name	Load Order
world	city	Normal Priority

Tables: 3

Tables: 1

OK Cancel

- Apache

Kafka

Introduction

<https://kafka.apache.org/intro.html>

- Qlik Replicate - Ingesta y replicación universal de datos

<https://www.qlik.com/es-es/products/qlik-replicate>

- What is Confluent Platform?

<https://docs.confluent.io/current/platform.html>

- Introducing the Kaazing.io Kafka Sandbox

<https://kaazing.com/introducing-the-kaazing-io-kafka-sandbox/>



Morris & Opazo



Advanced
Consulting
Partner

Data & Analytics
Competency

Solution Provider

Public Sector Partner

Immersion Day Partner

Expertos en Tecnología de la Nube

We Love the Cloud

www.morrisopazo.com

contacto@morrisopazo.com