

Datalake

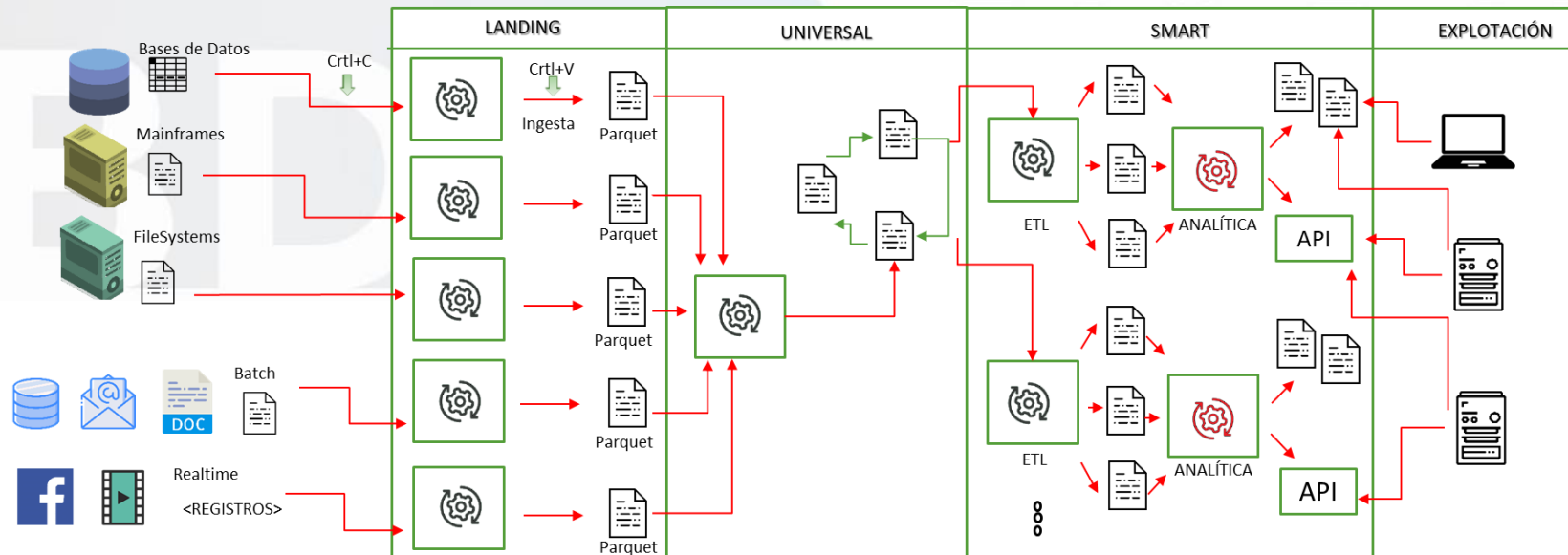
PARTE DE LA FORMACIÓN BIG DATA ENGINEER DE BIG DATA
ACADEMY

Concepto

BIG DATA
ACADEMY

Datalake

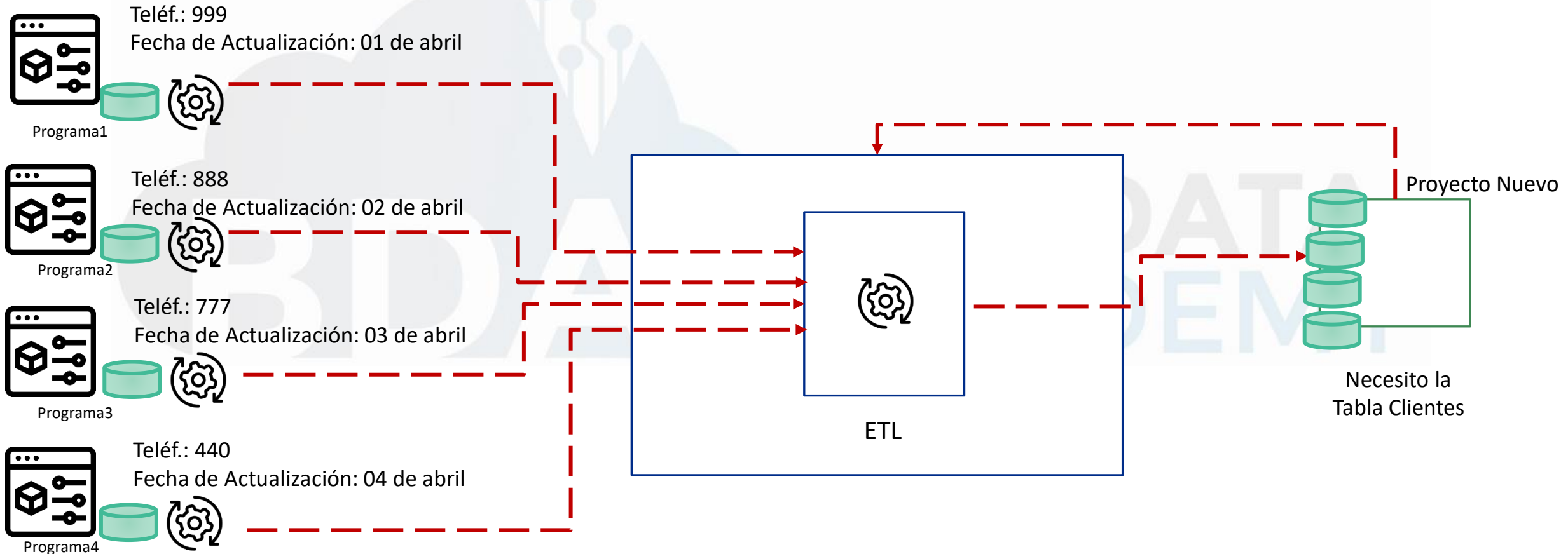
Es un gran repositorio de almacenamiento de datos en bruto, sin importar su estructura, que tiene como **objetivo ejecutar procesos de analítica**.



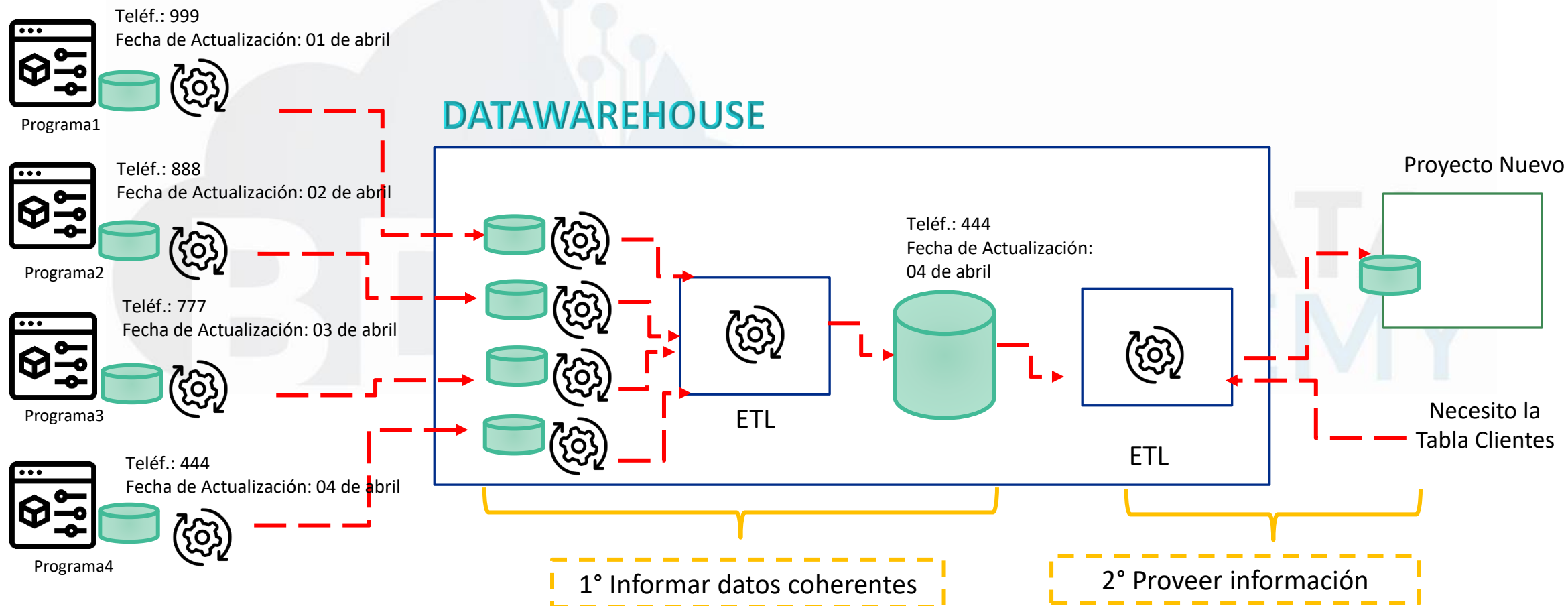
¿Por qué un Datalake?

Antes de un Datawarehouse

ISLAS DE FUENTES DE INFORMACIÓN

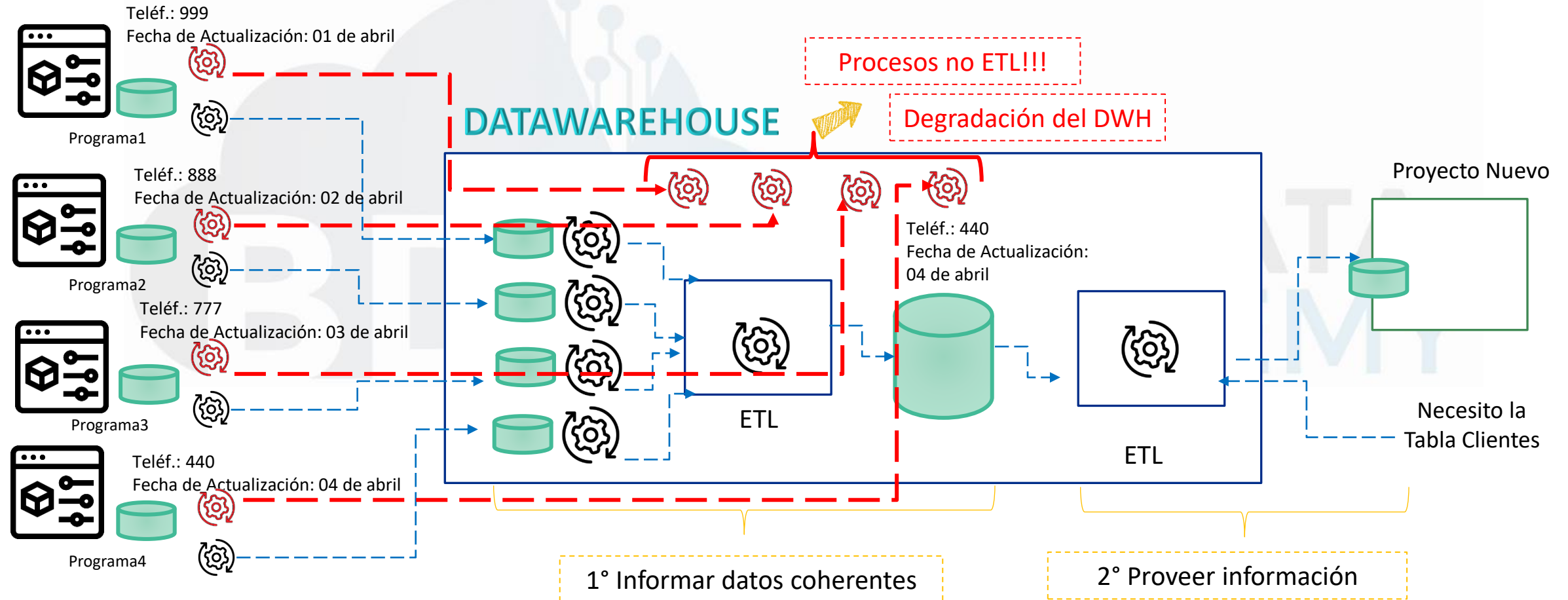


Con un Datawarehouse

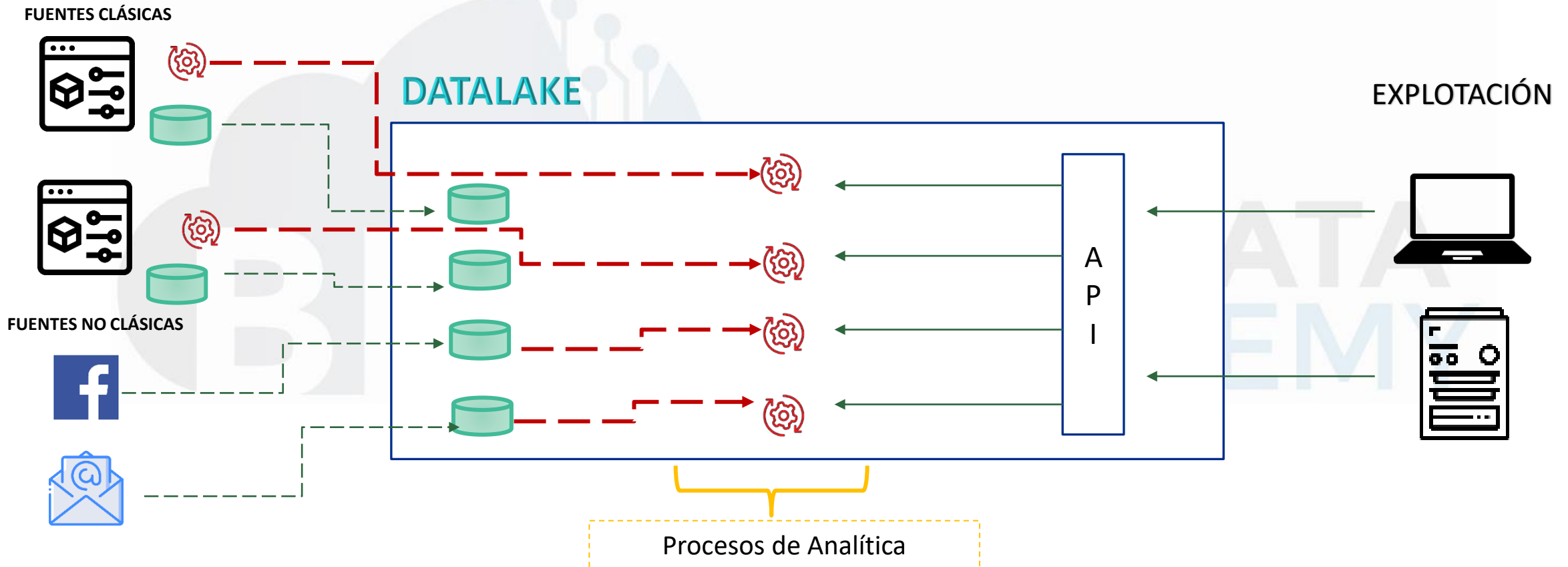


¿Analítica sobre Datawarehouse?

ISLAS DE FUENTES DE INFORMACIÓN



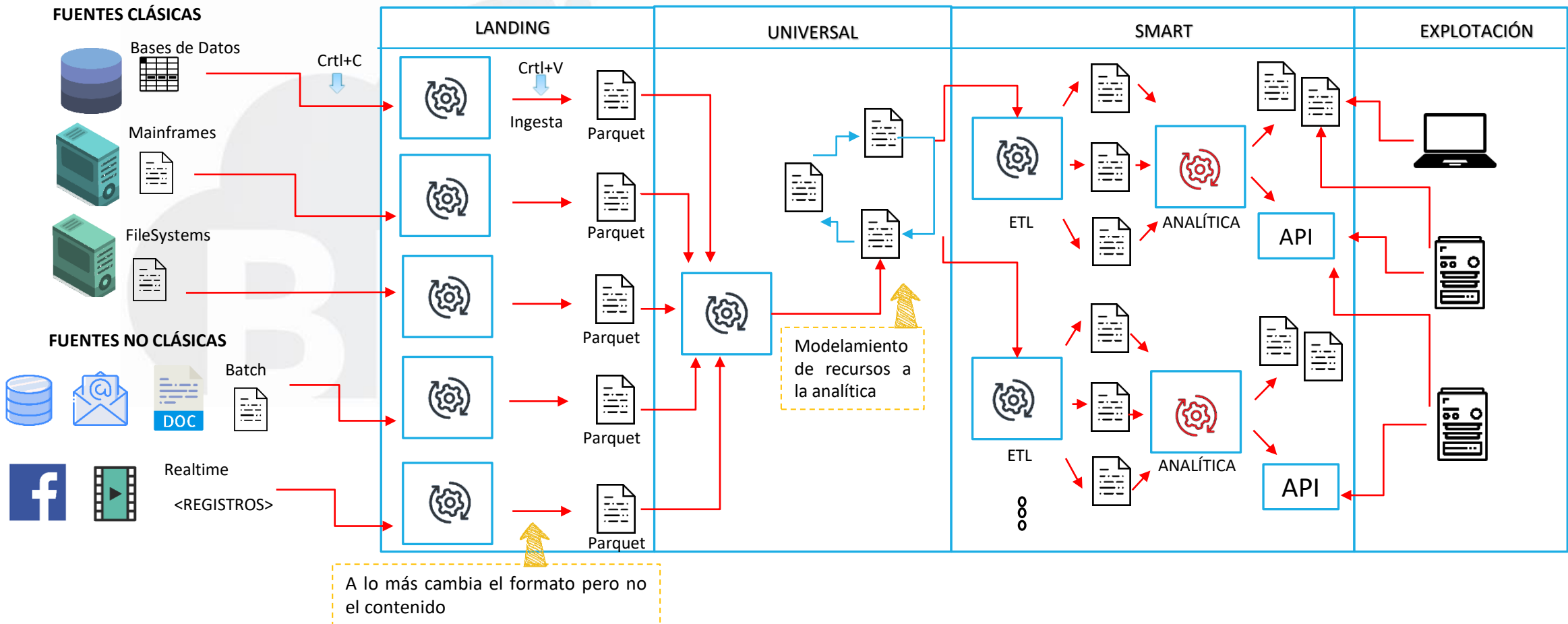
Analítica sobre el Datalake



Arquitectura de un Datalake

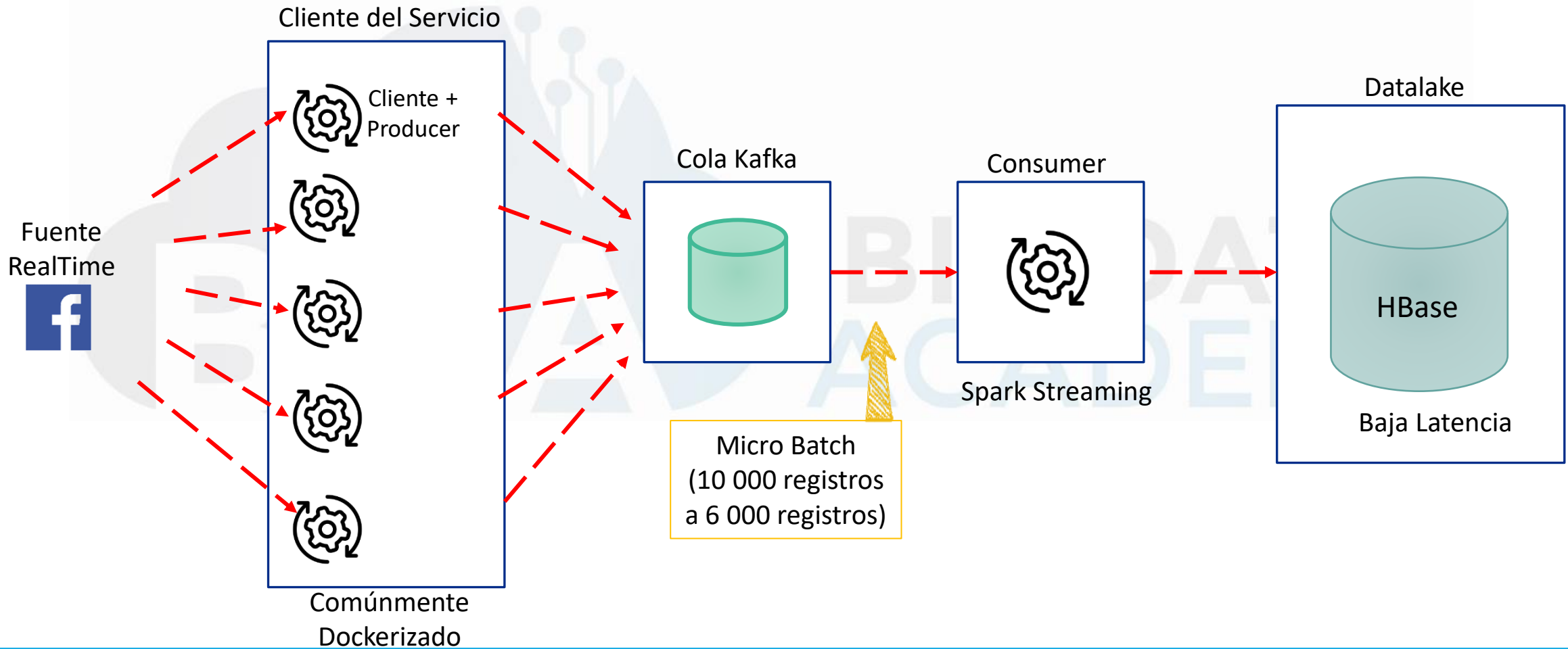


Arquitectura de un Datalake



Un ejemplo de patrón sobre el Datalake

Patrón de ingesta real time



Ejercicios teóricos

Ejercicios teóricos

1. ¿Qué es un Datalake?
2. Explique la capa LANDING
3. Explique la capa UNIVERSAL
4. Explique la capa SMART
5. ¿Por qué es necesario usar el patrón de ingesta para real time?

Resumen

Hablemos...

