

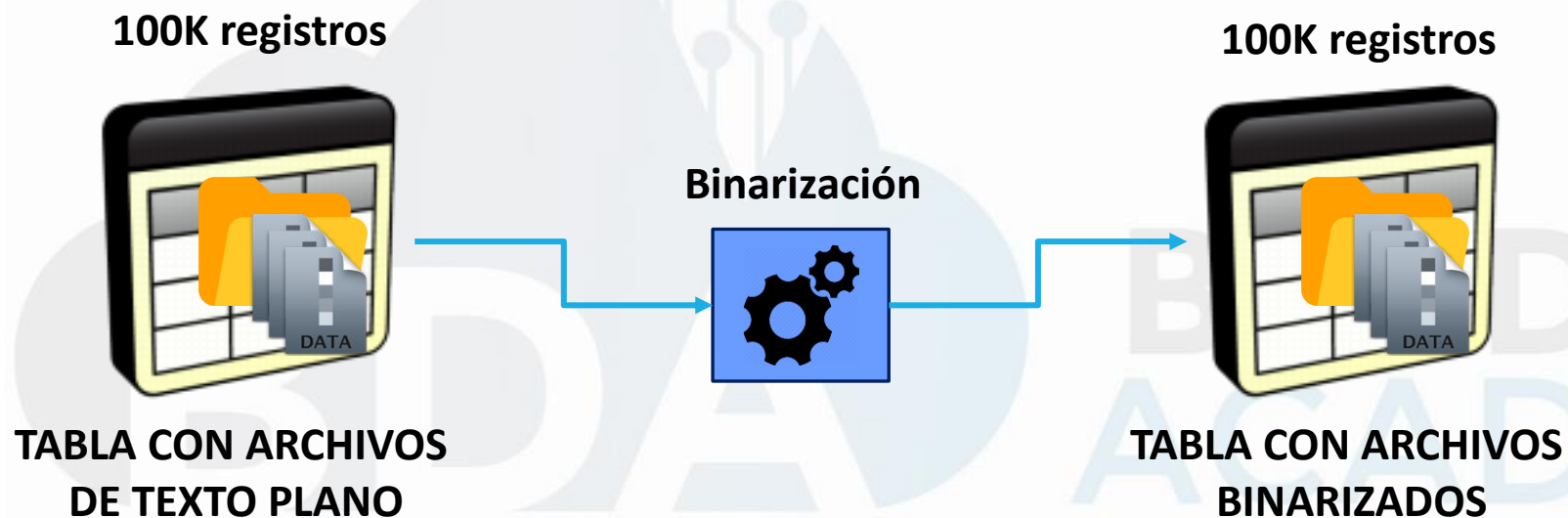
---

# Laboratorio 07

---

HIVE – FORMATOS BINARIZADOS

# Los formatos binarios de rápido procesamiento



El procesamiento sobre tablas binarizadas es hasta 10 veces más rápido

El proceso de binarización transforma los archivos de texto plano de una tabla, y coloca los archivos binarizados en otra tabla.

# Formatos binarizados



- Es el formato más rápido de todos
- Soporta datos estructurados y semi-estructurados
- No permite actualizaciones de los registros
- No permite usar tipos de dato DATE



- Es un fork (derivado) del formato PARQUET
- Soporta datos estructurados y semi-estructurados
- **Sí permite actualizaciones de los registros**
- Sí permite usar tipos de dato DATE

**¿Cuál usar?, siempre que se pueda, usar el formato DELTA**, sin embargo en algunos ecosistemas no hay compatibilidad con DELTA, en ese caso usar PARQUET.