



Indice

- Introducción
- Extracción
- Transformación
- **Carga**



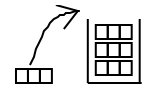
Indice

- Introducción
- Extracción
- Transformación
- Carga

Carga

Planteamiento general

- Carga (**loading**) lleva los datos al DW.
- Carga puede necesitar mucho tiempo:
 - Planificar → intentar automatizar todos los procesos
- Carga inicial mueve grandes volúmenes.
- Cargas posteriores mueven volumen de datos más pequeño.
- El “negocio” determinar el ciclo de las cargas.





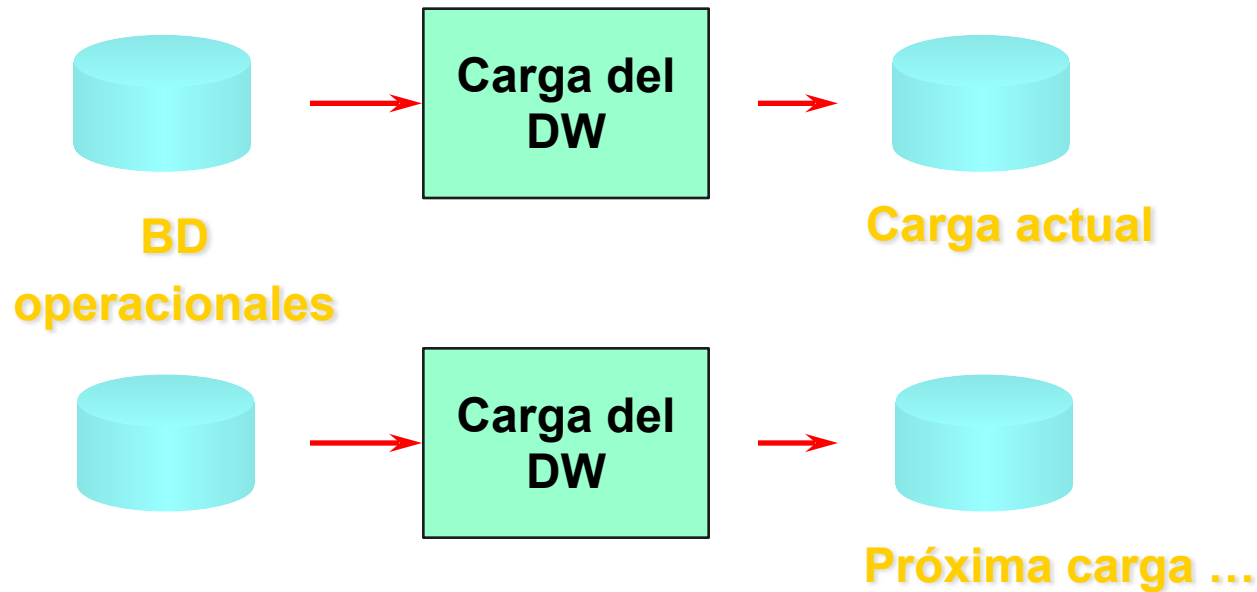
Carga

Capturando los cambios de datos para refrescar

- Capturar nuevos datos de hechos
- Capturar datos de dimensión cambiados
- Métodos
 - Remplazar datos a gran escala
 - Comparar instancias de bases de datos
 - Comprobar/escanear tiempo (Time stamping)
 - Triggers en bases de datos
 - Logs en bases de datos
- Técnicas híbridas

Carga

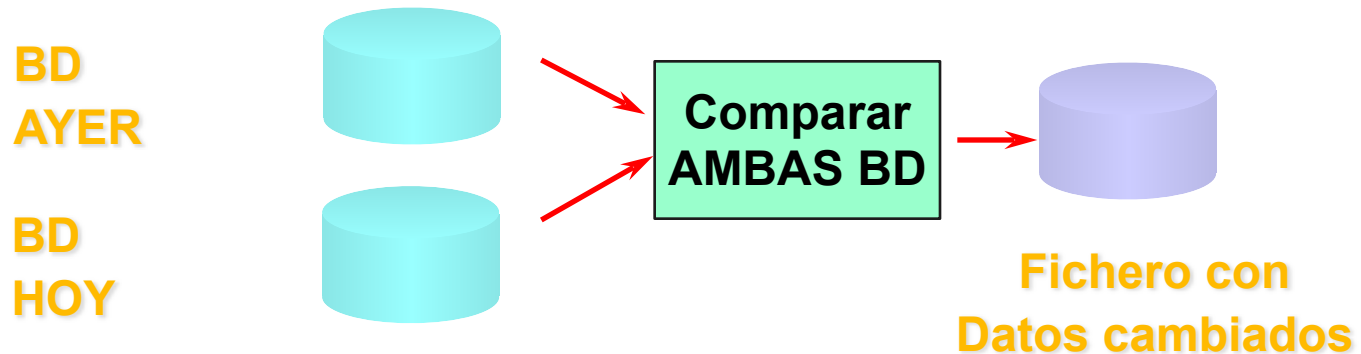
Cambios para refrescar. Remplazar a gran escala.



- Costosa en tiempo

Carga

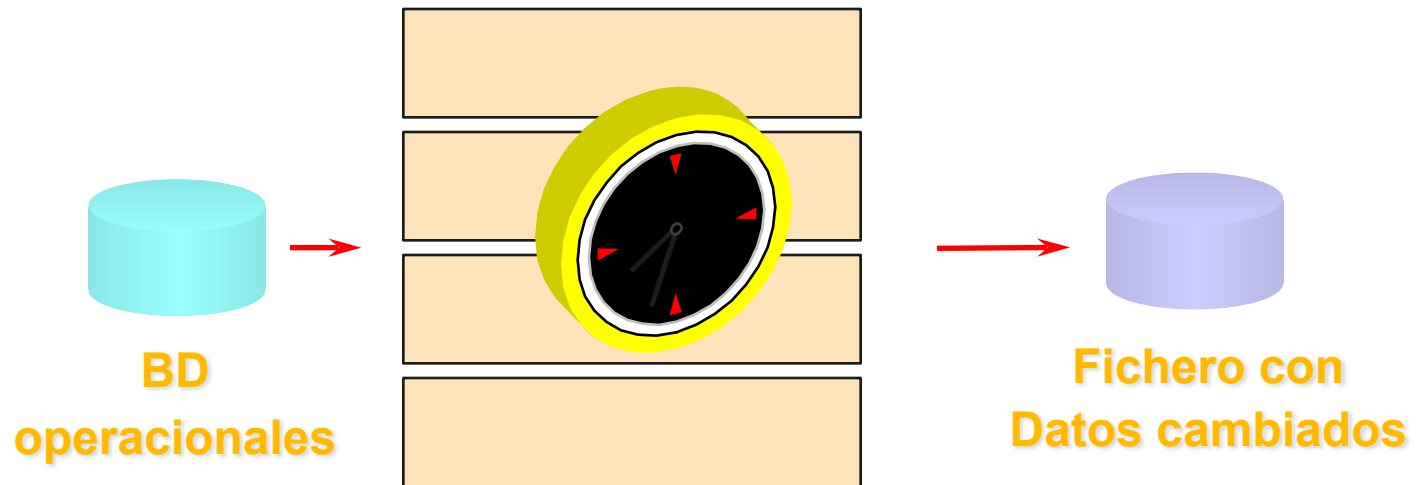
Cambios para refrescar. Comparar instancias de BD.



- Simple pero todavía cara
- Fichero con cambios
 - Cambios de datos operacionales desde último refresco

Carga

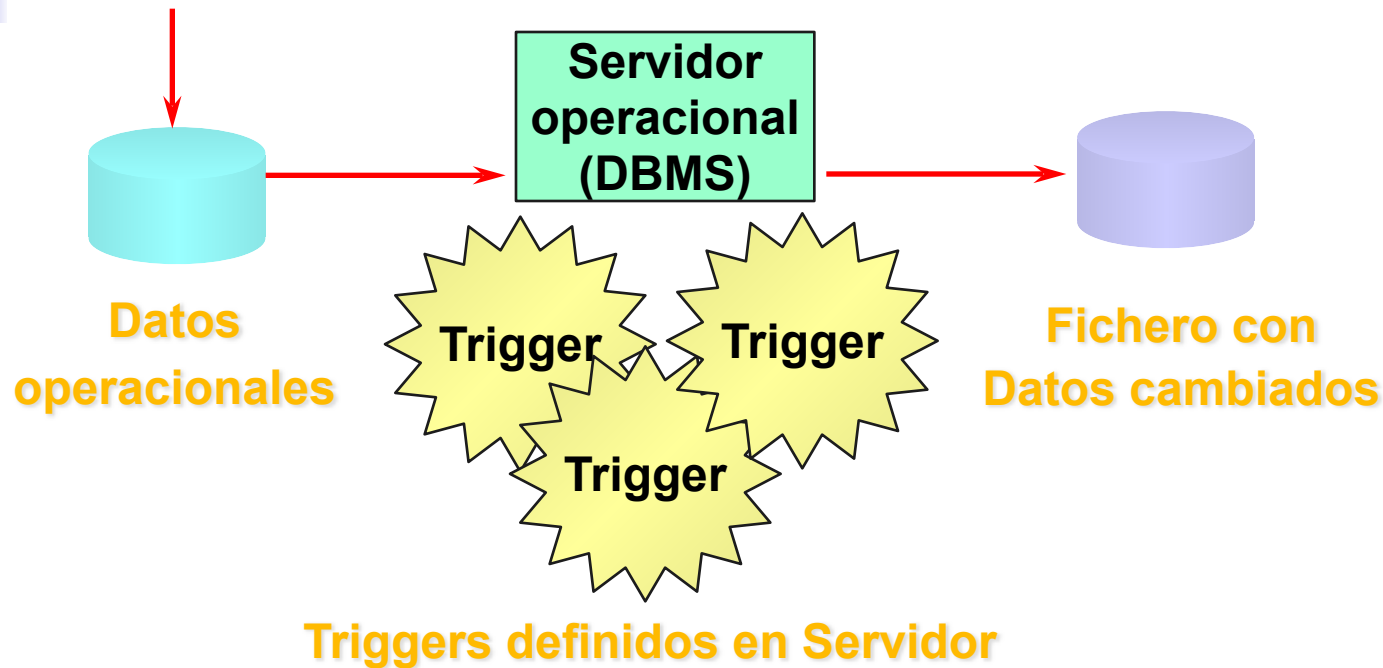
Cambios para refrescar. Time y Date stamping



- Rápida comprobación para los registros cambiados desde última extracción
- Fichero actualizado respecto de fechas
- No detecta los datos borrados

Carga

Cambios para refrescar. Triggers en BD



- Datos cambiados interceptados a nivel de servidor
- Necesidades extras de mantenimiento



Carga

Cambios para refrescar. Entonces QUE ???

- Analizar cada método de forma individual
- Considerar una solución híbrida si un solo método no es adecuado



Procesos ETL

(Extracción, Transformación y Carga)

Fin Tema 5