

Notions BI

Équipe BI

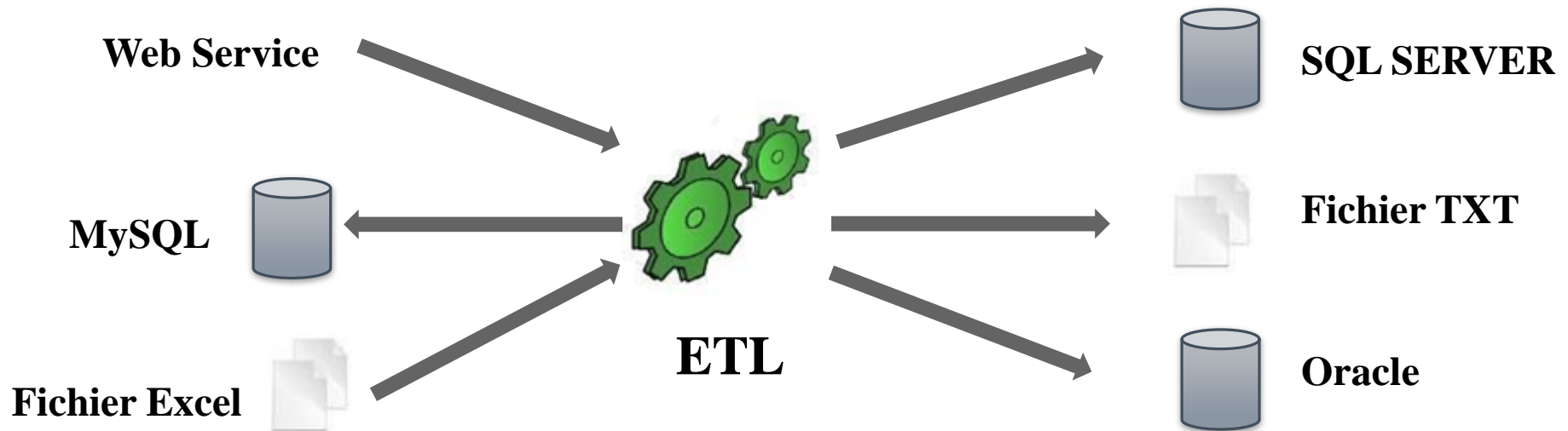
Plan

1. ETL : Extract-Transform-Load
 2. ODS (Operational Data Store) et SA (Staging Area)
 3. Architecture globale d'un système décisionnel
 4. Méthodologie (Cycle de vie dimensionnel)
-

ETL : Extract-Transform-Load

ETL est une technologie intergicielle (Middleware) permettant de :

- Effectuer des synchronisations massives des données d'une banque de données vers une autre.
- Extraire, transformer, nettoyer et charger les données entre différents sources de données hétérogènes.



ETL : Extract-Transform-Load

Propriétaire

- SSIS
- ODI (Oracle Data Integrator)
- Feature Manipulation Engine
- IBM InfoSphere DataStage
- Oxio Data Intelligence Solution

ETL

- Statistica ETL
- Talend

Open Source

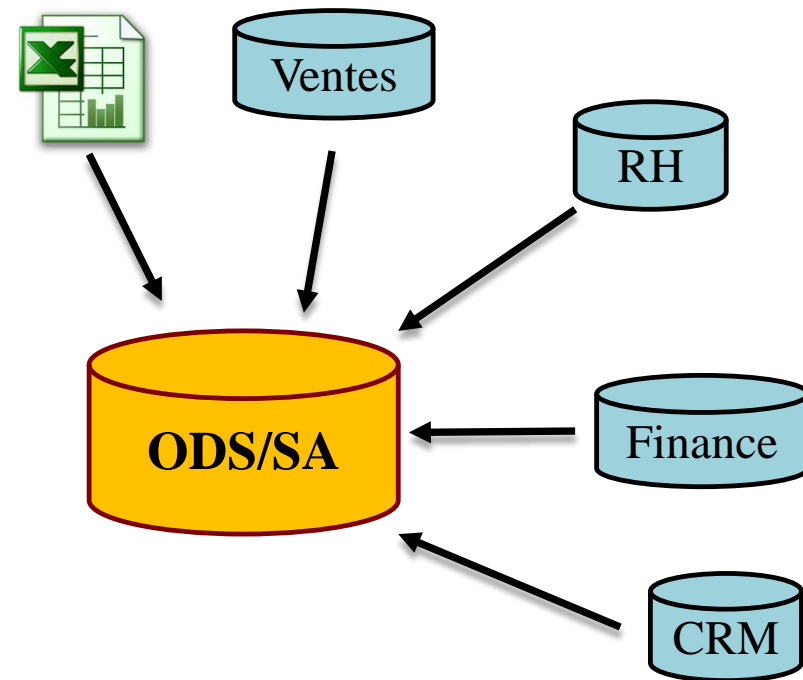
- Pentaho Data Integration
- Talend
- CloverETL
- Scriptella
- GeoKettle

ODS (Operational Data Store) et SA (Staging Area)

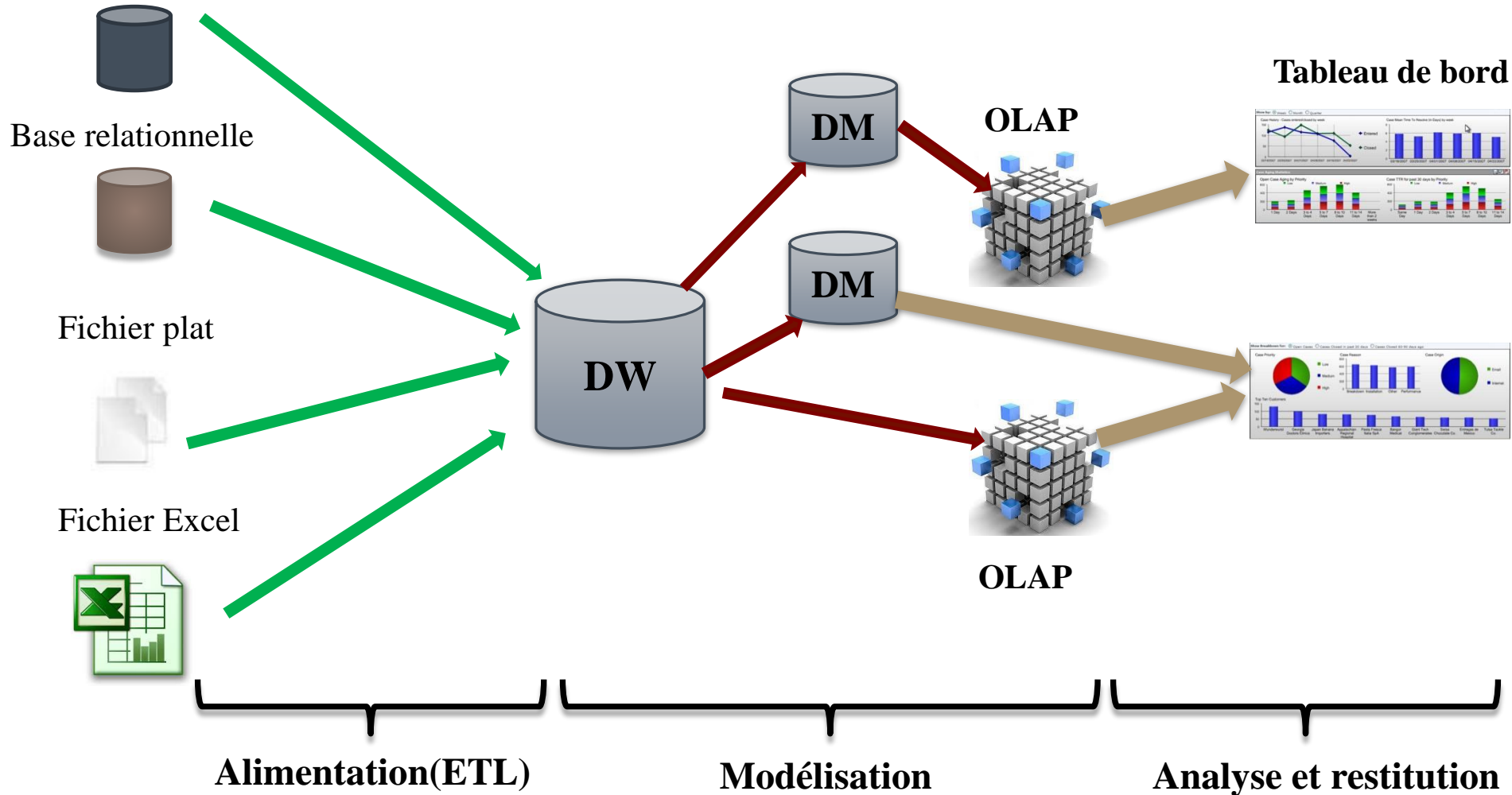
- C'est une base de données conçue pour centraliser les données issues de sources hétérogènes. Le modèle d'une base des données (ODS ou SA) est un modèle relationnel classique identique au modèle du système de production.

Leurs différences :

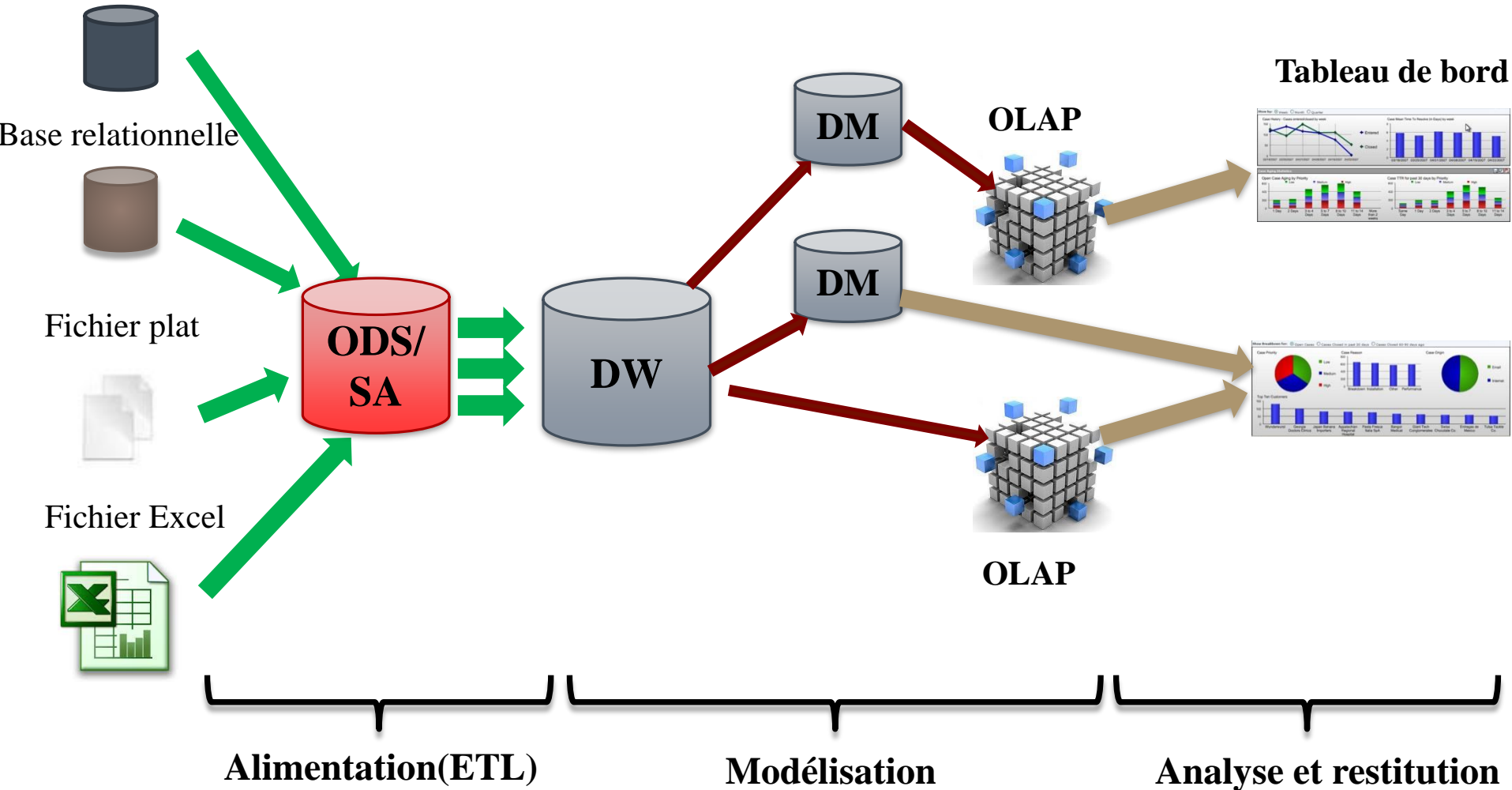
Dans le cas du **staging Area**, les données sont détruites directement après avoir été chargées dans le DataWarehouse mais pas pour l'**ODS** où les données auront quand même une durée de vie plus longue.



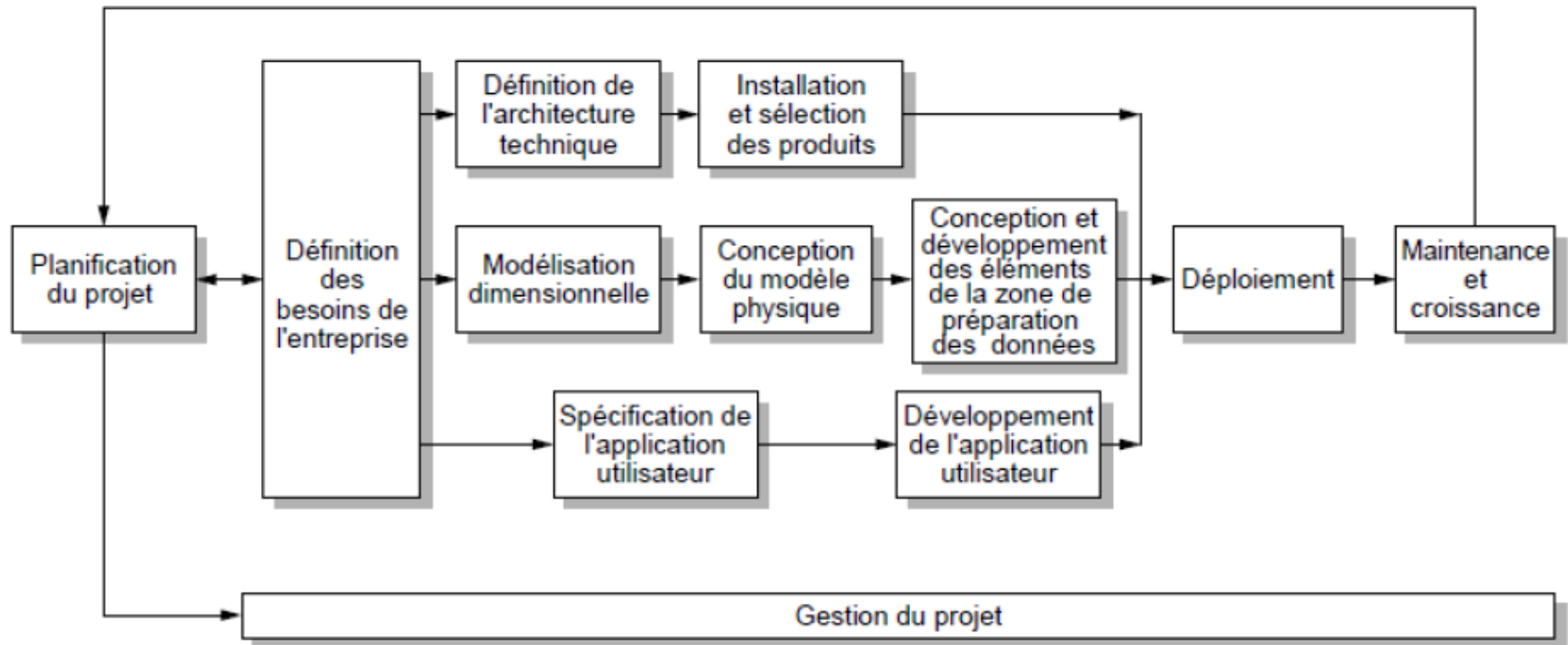
Architecture globale d'un système décisionnel



Architecture globale d'un système décisionnel



Méthodologie (Cycle de vie dimensionnel)



Méthodologie (Cycle de vie dimensionnel)

- **Planification du projet :**

la définition et l'étendue du projet décisionnel et elle se concentre sur les besoins en termes de ressources ainsi l'affectation des tâches, leurs durées et leur séquencement.

- **Définition des besoins de l'entreprise :**

les facteurs qui conduisent l'entreprise à vouloir définir ses besoins et les traduire pour les intégrer lors de la phase de conception. Ces besoins sont le point de départ de trois tâches en parallèle qui sont la technologie, les données et les interfaces utilisateur.

- **Définition de l'architecture technique :**

prendre en considération les besoins, l'environnement technique existant et l'orientation technique planifiée afin de choisir la structure de l'architecture technique

Méthodologie (Cycle de vie dimensionnel)

- **Modélisation multidimensionnelle :**

une analyse plus détaillée des données des systèmes opérationnels et en la couplant avec les besoins dégagés dans la phase précédente nous développons le modèle dimensionnel

- **Conception du modèle physique :**

les structures physiques nécessaires pour l'implémentation de la base de données logique

- **conception et développement des éléments de la zone de préparation des données :**

Cette tâche se déroule en trois étapes majeures Extract-Load-Transform.

Méthodologie (Cycle de vie dimensionnel)

- **Spécification de l'application utilisateur :**

les maquettes d'états, les critères de l'utilisateur et les calculs nécessaires. Cette tâche garantit que l'équipe de développement et les utilisateurs ont intégré une compréhension commune des applications qui vont être livrées.

- **Déploiement :**

le point de convergence de la technologie, des données et des applications utilisateur. Dans cette tâche, il faut mettre en place les processus de communication ainsi prendre en compte les demandes d'évolution ou de correction.

- **Gestion du projet :**

Cette garantit que les activités du cycle de vie dimensionnel restent sur la bonne voie et sont bien synchronisées et elle contrôle l'état d'avancement du projet, la détection et la résolution des problèmes.