

# Document – Schéma OLAP

## 1. Objectif du schéma OLAP

Le schéma OLAP a pour objectif de permettre l'analyse des transactions Stripe à grande échelle. Il est utilisé pour :

- analyser les revenus
- suivre les performances des marchands
- étudier le comportement des clients
- détecter et analyser la fraude
- produire des rapports de conformité

Ce schéma est séparé du système OLTP afin de ne pas impacter les performances transactionnelles.

## 2. Type de schéma et grain des données

Le modèle OLAP est conçu sous la forme d'un schéma en flocon (snowflake).

La table centrale **FACT\_TRANSACTION** contient les événements transactionnels.

Chaque ligne de cette table correspond à une transaction unique.

Les dimensions décrivent les axes d'analyse autour de cette transaction :

- temps
- client
- marchand
- produit
- localisation
- appareil
- fraude
- conformité

## 3. Analyse des séries temporelles

Le schéma inclut une dimension temps détaillée (DIM\_DATE) permettant :

- des analyses quotidiennes
- des comparaisons hebdomadaires, mensuelles et annuelles
- des analyses de tendances (évolution des revenus, taux de fraude)

Les hiérarchies de temps (jour → mois → année, jour → semaine) facilitent les analyses temporelles.

#### **4. Gestion des jointures à grande échelle**

Les jointures entre la table de faits et les dimensions sont optimisées grâce à :

- l'utilisation de clés substituts numériques
- une structure simple de jointures entre FACT\_TRANSACTION et les dimensions

Pour améliorer les performances :

- la table FACT\_TRANSACTION peut être partitionnée par date
- seules les partitions nécessaires sont lues lors des requêtes analytiques

Cette approche permet de traiter de grands volumes de données tout en conservant de bonnes performances.

#### **5. Pré-agrégations et optimisation des requêtes**

Afin de réduire le temps de réponse pour les requêtes les plus fréquentes, des pré-agrégations peuvent être mises en place.

Exemples de pré-agrégations :

- revenus journaliers par marchand
- revenus mensuels par pays
- taux de fraude par période

Ces agrégations peuvent être stockées sous forme de tables de synthèse.

Cela permet d'éviter de recalculer les mêmes agrégats à chaque requête.

#### **6. Optimisation des performances**

Plusieurs stratégies simples sont utilisées pour garantir de bonnes performances analytiques :

- partitionnement de la table de faits par date
- modèles incrémentaux pour charger uniquement les nouvelles données
- utilisation du cache des outils de visualisation (BI)

Ces optimisations permettent une analyse rapide sans complexité excessive.

## **7. Lien avec le système OLTP => source de vérité**

Les données OLAP sont alimentées à partir du système OLTP via des pipelines de données.

Les données sont :

- extraites
- transformées
- chargées dans le Data Warehouse

Le système OLTP reste la source de vérité, tandis que le système OLAP est dédié à l'analyse.

## **8. Conclusion**

Ce schéma OLAP permet :

- une analyse efficace des données transactionnelles
- une bonne évolutivité face à l'augmentation des volumes
- une séparation claire entre traitement transactionnel et analytique

L'architecture proposée répond aux besoins d'analyse avancée tout en restant simple et compréhensible.