

# Environnement:

Ce document vous présente l'architecture de la plateforme de données d'une société de commerce électronique nommée SoftCart.

SoftCart utilise une architecture hybride, avec certaines de ses bases de données sur site et d'autres sur le cloud.

## Outils et technologies :

- Base de données OLTP - MySQL
- Base de données NoSql - MongoDB
- Entrepôt de données de production – DB2 sur Cloud
- Entrepôt de données de préparation – PostgreSQL
- Plateforme Big Data - Hadoop
- Plateforme d'analyse de Big Data – Spark
- Tableau de bord de Business Intelligence - IBM Cognos Analytics
- Pipelines de données - Apache Airflow

## Processus:

La présence en ligne de SoftCart se fait principalement via son site Web, auquel les clients accèdent à l'aide de divers appareils tels que des ordinateurs portables, des téléphones portables et des tablettes.

Toutes les données du catalogue des produits sont stockées sur le serveur MongoDB NoSQL.

Toutes les données transactionnelles telles que l'inventaire et les ventes sont stockées dans le serveur de base de données MySQL.

Le serveur Web de SoftCart est entièrement piloté par ces deux bases de données.

Les données sont périodiquement extraites de ces deux bases de données et placées dans l'entrepôt de données intermédiaire exécuté sur PostgreSQL.

L'entrepôt de données de production se trouve sur l'instance cloud du serveur IBM DB2.

Les équipes BI se connectent à IBM DB2 pour créer des tableaux de bord opérationnels. IBM Cognos Analytics est utilisé pour créer des tableaux de bord.

SoftCart utilise le cluster Hadoop comme plate-forme Big Data où toutes les données sont collectées à des fins d'analyse.

Spark est utilisé pour analyser les données sur le cluster Hadoop.

Pour déplacer des données entre OLTP, NoSQL et l'entrepôt de données, des pipelines ETL sont utilisés et ceux-ci s'exécutent sur Apache Airflow.