

# Processing

MOC: [\\$INFORMATIQUE \\$DATA ENGINEERING](#)

Source :

Date : 2024-02-24

---

Composante distribué du [Big data](#) utilisé dans le traitement conséquent de donnée. C'est l'ensemble des opération faites sur les données pour les analyser, transformer ou extraire des informations utiles. Implique souvent des algorithmes de type [MapReduce](#).

Les deux manières de faire du processing est le [batch processing](#) et le [stream processing](#)

**batch processing** :: je prend un gros ensemble de donnée connue et je le transforme. Cette méthode utilise des solutions de [Storage](#). ([SPARK](#) or [Hadoop](#))

**stream processing** ::: traite plein de **petites données** rapidement grace au [Stream](#). peut choisir d'être rapide ou lente selon la demande (FLINK leader)