# Hadoop

## La solution Hadoop

### A

Hadoop : système qui est apparu pour résoudre aux problèmes des 3V.

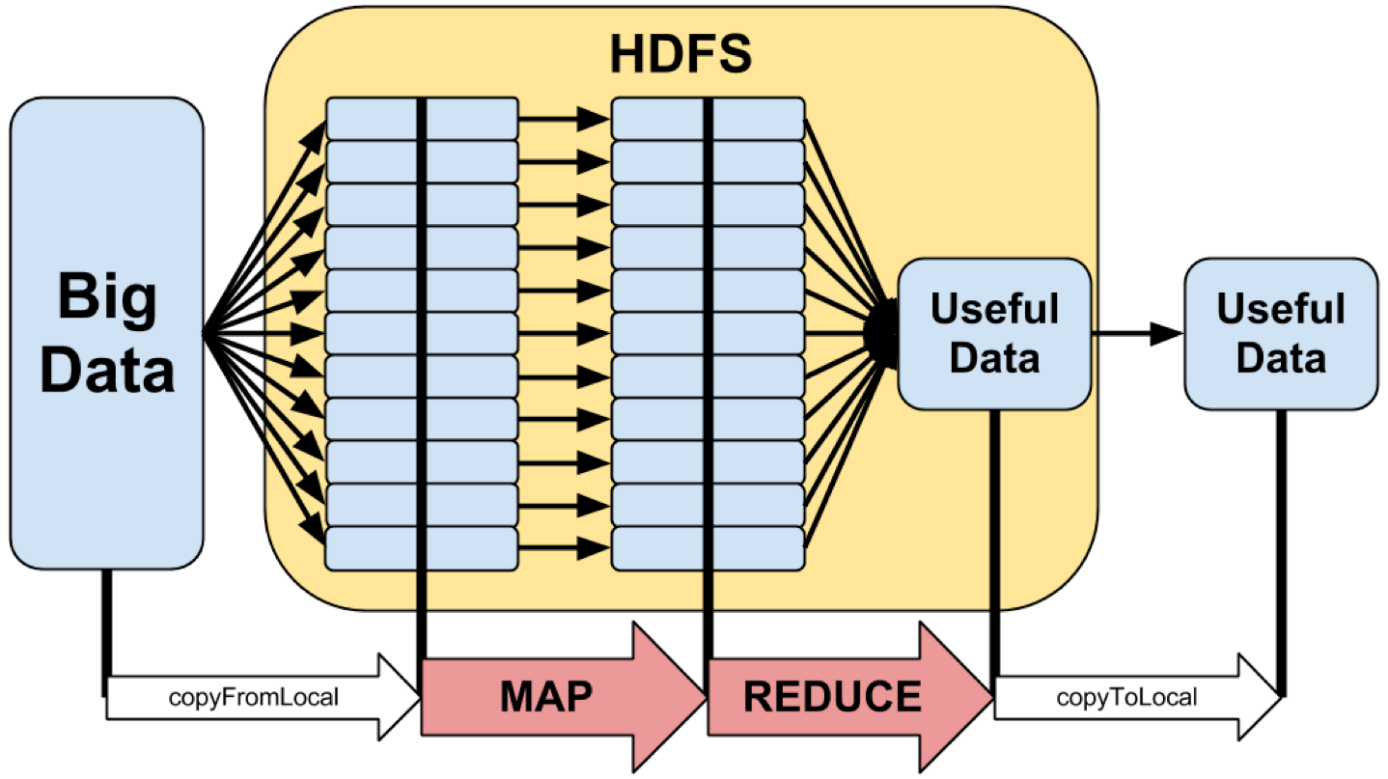
Hadoop est un framework libre et open source écrit en Java destiné à faciliter la création d’applications distribuées (au niveau du stockage des données et de leur traitement) et échelonnables (scalables) permettant aux applications de travailler avec des milliers de nœuds et des pétaoctets de données.

### Avantages

* Gestion de gros volumes de données
* Efficacité de stockage
* Bonne capacité de récupération de données
* Évolutivité horizontale
* Moindre cout

### A

Le noyau d’Hadoop est constitué d’une partie de stockage : HDFS (Hadoop Distributed File System), et d’une partie de traitement appelée MapReduce



*Hadoop fractionne les fichiers en gros blocs et les distribue à travers les nœuds du cluster. Pour traiter les données, il transfère le code à chaque nœud et chaque nœud traite les données dont il dispose. Cela permet de traiter l'ensemble des données plus rapidement et plus efficacement que dans une architecture supercalculateur plus classique.*

### A

A

•

—

## A

A

•

—