

# Hadoop Distributed File System (HDFS)

Data science  
Master 2 ISIDIS

SÉBASTIEN VEREL

verel@lisic.univ-littoral.fr

<http://www-lisic.univ-littoral.fr/~verel>

Université du Littoral Côte d'Opale  
Laboratoire LISIC  
Equipe OSMOSE

# Hadoop Distributed File System (HDFS)

## HDFS est :

- un **système de fichier** (file system) au même titre que fat32, ext3FS, NTFS, etc.
- un système de fichier **distribué** comme NFS (network file system)

## Buts de ce File System

- Stocker une grande quantité de donnée (fichiers)
- Donner accès aux données à de nombreux clients distribués sur le réseau

## Remarque :

Projet de la fondation Apache, Open Source.

# Support de cours

- Bilel Derbel,  
Université de Lille 1  
Programmation Parallèle et Distribuée,  
<https://sites.google.com/site/bilelderbelpro/home/teaching>
- Benjamin Renaut,  
MBDS université de Nice Sophia Antipolis  
Cours big data,  
<http://cours.tokidev.fr/bigdata/>