

Relationnel NS SQL  
Oracle  
Microsoft => SSRS

BI → DW / cubes

Big Data Hadoop

Parquet → column store

HBase → Database

HDFS → File Store

Spark / Scala

BI  
BI

NoSQL

→ Big Data

- MongoDB → Big Compute

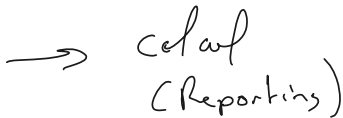
- Cassandra (column store)

- Redis (cache def / Vateur)

# Data / Compute



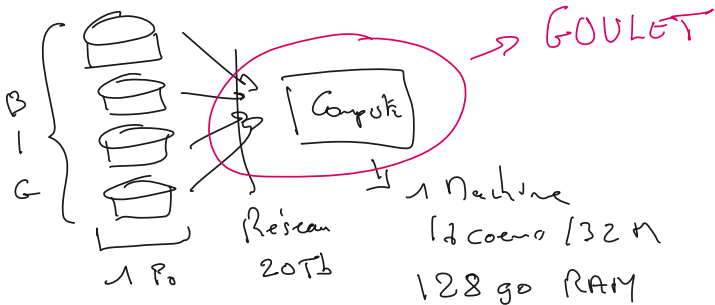
BDP



calcul  
(Reportings)

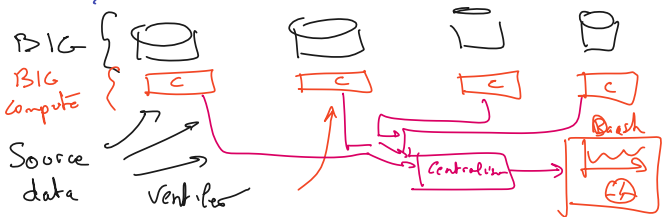
---

# Big Data / Small Compute



---

# BIG Data / Big Compute



# Cluster Hadoop



client

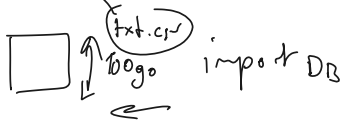
machine

```
> hdfs cp ... /hdfs/ txt.txt
```

sparks

→ Analytique  
Scoba

aggrégation



## MA Doo D.

File system : HDFS

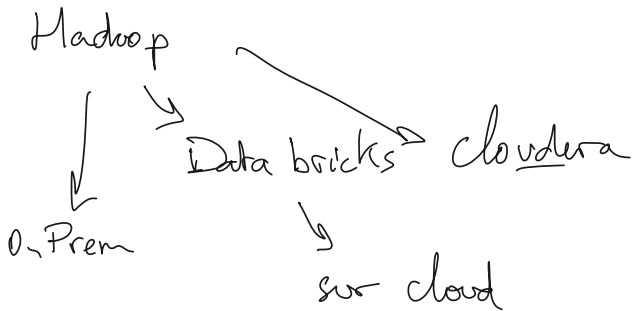
Base DD : HBase

CSV → vue → Parquet

Analytique → Spark

Message Queue:

Kafka



"Traditionnel" ↔ Hadoop

Coherence:

voir théorème de  
CAP