

PUBLICATION INTERNE

Les plateformes Big Data

AMRANE Abdesalam

Mai 2017

03 Rue des Frères Aïssou – Ben Aknoun – ALGER – ALGERIE
Tél. : 021 91 62 05 à 08 – Fax : 021 91 21 26
[http : //www.cerist.dz](http://www.cerist.dz)

PUBLICATION INTERNE

Les plateformes Big Data

AMRANE Abdesalam

Mai 2017

03 Rue des Frères Aïssou – Ben Aknoun – ALGER – ALGERIE
Tél. : 021 91 62 05 à 08 – Fax : 021 91 21 26
[http : //www.cerist.dz](http://www.cerist.dz)

AMRANE Abdesalam

Les plateformes Big Data

ISRN CERIST- DSISM/RR--17-000000008--DZ

Nom de l'organisme responsable CERIST

Date de publication Mai 2017

Résumé: Avec la croissance exponentielle du volume de données et la nécessité de l'analyse de ces données, la conception et la mise en œuvre d'une plateforme d'analyse de données Big Data pour gérer, stocker et extraire des prévisions de toutes les données numériques devient primordiale. Une variété de plateformes de Big Data existe, nous avons étudié certaines d'entre eux et nous avons présenté un tableau comparatif avec différentes caractéristiques. La plateforme qui nous semble très intéressante côté performance et traitement temps réel est Spark.

Mots clés: BigData, Hadoop, Spark, Flink, Storm, MapReduce

Sommaire

1. Introduction	6
2. Les approches de construction de plateformes Big Data	7
3. Les plateformes open-source	7
3.1. Apache Hadoop	8
3.2. Apache Spark.....	10
3.3. Apache Flink.....	12
3.3.1. Modèle de traitement des données.....	12
3.3.2. Modèle de traitement par lots	13
3.4. Apache Storm	14
4. Comparaison des plateformes.....	16