



# Étude de cas sur le traitement de données (Big Data)

## Migration de traitements Spark vers AWS

### À propos de Digiteka



Digiteka est une société développant de nombreux produits axés sur la vidéo en ligne. L'entreprise se positionne sur le marché du B2B et a comme client de nombreux médias tel que Euronews, Reuters, Non-Stop People, Sport 365 ou l'INA. Une part importante de l'activité de Digiteka est l'analyse du comportement des utilisateurs de son lecteur vidéo. Les données récupérées sont conséquentes (80 GB non compressés par jours) et sont traités quotidiennement chez un hébergeur tiers (OVH). Digiteka souhaite migrer ce service critique vers AWS afin de bénéficier d'une infrastructure plus résiliente.

### LE DÉFI

Le traitement de ces données s'effectue actuellement sur des clusters Spark Mesos et s'effectue quotidiennement en 5 heures. La migration et l'exécution du service ne doit occasionner aucune fuite de données et du fait de la criticité du service, se doit d'être résilient. La solution doit également avoir un coût maîtrisé.

### LA SOLUTION

S3 est mobilisé comme data lake pour les données brutes et transformées. EMR est utilisé pour les traitements des données Spark en conjonction de AWS Glue. Ce dernier prépare les données afin d'être lues par Amazon Athena. Les données analysées sont traduites sous la forme de dashboard grâce à Quicksight. L'orchestration des services et leurs actions sont assurés par Step Functions.

### LES AVANTAGES

#### Performance

Des traitements bien plus rapides abaissant la durée de 5 heures à presque 2 heures 30 (deux fois plus rapide).

#### Financier

Du fait d'une performance accrue et d'un paiement à la ressource consommée, plus de 82% d'économie sont réalisées.

#### Résilience

Grâce aux services managés d'AWS et aux engagements de disponibilité, le maintien en service est assuré.

### A propos de COREXP ET D'AWS

Agréé Advanced Consulting Partner et Value-Added Reseller, Corexpert possède plus de 13 années d'expérience dont plus de 9 dans le cloud AWS. La société a rejoint le groupe TeamWork en 2017 et déploie ses compétences sur plus de 15 localisations dans le monde avec une forte expertise dans SAP, le BigData, les Web stacks et accompagne les migrations de toute taille sur AWS.

### ETAPES SUIVANTES

Pour en savoir plus sur la façon dont AWS et Corexpert peuvent vous aider, visitez [corexpert.net/fr](http://corexpert.net/fr)