

Maxime GILLOT

Nantes, France.

06.33.65.91.86

gillotmaxime@yahoo.com

[Linkedin](#)

[Mon site internet](#)

30 ans

Permis B

**Data Engineer
Développeur Big Data
Expert Spark****PROJETS PERSONNELS**

- Formateur occasionnel Apache kafka pour [Spark formation](#) (intra et inter entreprise)
- Création de formation sur [Apache Kafka](#) et [Apache Nifi](#) sur [Udemy](#) (1500 participants)
- POC de création de jeux vidéos, développement blockchain, site internet, Bot ...
la liste est disponible sur [mon site internet](#)

EXPÉRIENCE PROFESSIONNELLE

Crédit Agricole Technologie et Services (abrégé: CA-TS)

Data Engineer / Expert technique Spark

Janvier 2023 - Février 2024

Au sein de l'équipe socle data:

- Conception, Développement, Tester, Déployer et assurer le RUN et le support du facilitateur Spark interne à CA-TS.
 - Spark
 - Java
 - Yarn
 - BDD divers : MySQL, PostgreSQL, Teradata, Hbase, Hive, MSSQL, Mongo ...
 - Kafka
 - Api Rest
- Support build, run, architecture et conception aux différentes équipes de CA-TS (sur l'outil et sur la plateforme big data)
- Ordonnancement : Oozie, création de workflow
- Optimisation job Spark.
- Rédaction de la documentation technique, script xWiki.
- Formation, monter en compétences et bonne pratiques des différents collaborateurs sur la plate-forme big data.

Environnement technique :

Cloudera, Spark 2.4.4/3.2.2, Java 8/11, Hadoop, Kafka, Oozie, Linux, jenkins, Gitlab CI/CD, documentation Xwiki, Git, IntelliJ Méthode agile : Scrum

My Money Bank

Data Engineer AWS / Expert technique Spark

Juillet 2020 - Décembre 2022

Au sein de l'équipe Data publishing:

- Conception, Développement, Tester, Déployer et assurer le RUN de différents flux de données et applications afin de produire des reportings opérationnels.
 - Spark
 - Scala
 - Yarn
 - Postgresql

- Java 8
- AWS : S3, EMR, Athena, Glue
- Sharepoint
- Excel
- Documentation Antora

- Intégration continue : Gitlab CI/CD, airflow
- Ordonnancement: Airflow / Python , Création de DAG.
- Optimisation job Spark.
- Soutient aux différentes équipes Spark: optimisation de job, debug, bonne pratique de développement, tests unitaires.
- Participation active aux développement des librairies communes MMB.
- Création de rapport Excel avancé.
- Création de batch JAVA : Java / Spring.
- Rédaction de la documentation technique.
- Challenger le mapping fonctionnelles des données.

Environnement Technique :

AWS : S3, Athena, EMR, RDS, System Manager. Spark 2.6/3.1.1, scala/Java 8/11/14/17, gitlab Ci/CD, Kibana, Airflow, Sharepoint, Méthode agile : Scrum/Kanban

Sopra Steria

Data Engineer / référent technique Spark

Mai 2020 - Juillet 2020

Dans le cadre de la création d'un starter kit Spark à la CNAM :

- Concevoir, développer, documenter, tester des Flux Spark :
 - Spark
 - Hadoop
 - HBase
 - Java 8
 - Junit
- Développement dans un environnement sécurisé : Knox, Ranger, ranger KMS, Kerberos
- Dévops: jenkins, GitLab, artifactory
- Scripts linux/yarn/hadoop/hdfs : cycle de vie des applicatifs

Environnement technique :

Écosystème Hortonwork (Hadoop, Hdfs, YARN, Spark 2.3.2) Java 8, Linux, Jenkins, gitlab, Docker, IntelliJ, maven

Sopra Steria

Data Engineer

Mars 2018 - Avril 2020

Dans le cadre d'un POC, mise en place d'une filière Big Data à la CNAM (from scratch) :

- Concevoir, développer, documenter, tester et mettre en production différents flux pour alimenter un data lake:
 - Spark streaming
 - Batch Spark
 - Kafka stream, Kafka connect, custom SMT / Avro + schema registry
 - Flow Nifi, script Groovy, processors custom (Java 8)
 - Batch java/Hadoop
 - Montée de version applicative vers Hadoop 3 (Spark, Kafka, Hive)
- Dévops : Jenkins, Git/GitLab, Artifactory, Saltstack, Maven
- Scripts linux/yarn/hadoop/hdfs : cycle de vie des applicatifs
- Monitoring: création de dashboards Grafana (kafka,Yarn,nifi), configuration JMX, Prometheus, graphite exporter
- Scripts Hive, notebook Zeppelin (Livy, spark)

- Développement dans un environnement sécurisé : Knox, Ranger, ranger KMS, Kerberos
- Participation à la configuration des machines Spark, Hadoop, Nifi, YARN
- Mise en place d'un environnement de développement local (Docker, Hortonworks)
- Gouvernance de données avec Atlas, création linéage
- Méthode agile : SCRUM, iceScrum.
- Environnement de test : Junit5, hadoop-mini-cluster, Nifi mock framework.

Environnement technique :

Écosystème Hortonwork (Hadoop, Hdfs, YARN, Spark, Kafka (HDP et Confluent) , Nifi, Zookeeper, Hive, Ambari, ...), Avro, Java 8, Linux, Jenkins, gitlab, Docker, confluent, groovy, IntelliJ/Éclipse, maven, Sonar

Bouygues Telecom

Consultant Big Data

Octobre 2017 - Mars 2018

- Mise à jour et création de flux de données
 - Kafka stream
 - Flink
 - Hive
 - MongoDB
 - Oozie
 - Hdfs
 - Grafana
 - Java 8
 - Création de script Hql et Teradata pour la migration, création et mise à jour de table.
 - Monitoring: création de dashboards Grafana
 - Déploiement/build: Maven, Git/GitLab, SVN
 - Environnement de test: Junit 4
 - Étude d'impact
 - Méthode agile : SCRUM, Jira
- Environnement technique :
Écosystème Cloudera, Hadoop/hdfs, Hive, impala, MongoDB, Flink, oozie, Java 8, TeraData, git, svn

Sopra Steria

Développeur Big Data

Septembre 2017 - Mars 2017

Dans le cadre de mon stage de fin d'études:

- Concevoir, développer, documenter, tester différents batchs
 - Spark Batch et streaming
 - Cassandra
 - Hadoop (HDFS, YARN ...)
 - Hive
 - Java 8
 - Scala
 - Spring
- Translation de batch java/Spark vers Scala/Spark
- méthode agile : SCRUM, jira

Environnement technique :

Écosystème big data, (datastax/cloudera) Hadoop/Hdfs, Hive, Cassandra, git, java et scala, scrum.

FORMATION

2015 – 2017: Master recherche informatique – Université d'Angers

2011 – 2014: Licence scientifique mention INFORMATIQUE – Université d'Angers

COMPÉTENCES

Langue: Anglais professionnel

Big Data: Spark, Kafka, Yarn, Nifi, Hdfs, Hadoop, Nifi, Avro, Schema Registry, Hive, Java/Scala, Oozie ...

Cloud: AWS

Environnement: Linux, Windows, Git/GitLab, Eclipse/IntelliJ, Docker, Scrum

Devops: Jenkins, Maven, artifactory, Ansible, Gitlab CI/CD