



EMMANUEL HAREL

Data Engineer / DevOps

+15 années d'expérience

emmanuel.harel@gmail.com

FORMATIONS



DESS
Instrumentation
physique
Paris 12



Développeur Java
IFOCOP



Formation Kafka -
Confluent
Operations
Zenika



AWS Certified
Solutions
Architects
A Cloud Guru



Cloud
DevOps AWS
Blent AI



Data Engineer
with Python
DataCamp

EXPÉRIENCES

Germen (Garance)

mai 2023 - mai 2024

Développeur Backend Java / DevOps

Dans le cadre de la collaboration avec Garance, Germen crée des simulateurs pour la retraite, simulateur fiscal, pour les placements, et les projections des performances des clients.

Mes tâches effectuées :

- déploiement des applications
- développements backend de l'API REST en Java pour la refonte d'une application de gestion de l'épargne
- clean des données d'une base Postgres en Python
- développement from scratch d'une application en spring boot 3 pour l'enrichissement de données provenant de plusieurs datasources
- pagination multicritère pour les développeurs front et le data scientist

Equipe composée de :

- 1 Data Engineer / Backend Developer
- 4 Frontend développeurs
- 1 Data Scientist
- 1 DevOps
- 1 product owner
- 1 testeur

Compétences : Java 11/17, spring boot 2 et 3, swagger, REST API, hibernate, JOOQ, python, dataframe, spring security, microservices, OVH, docker swarm, métier de l'épargne, Algorithme de Monte-Carlo

Groupm

novembre 2019 - sept 2020

Java Backend développeur / Data Engineer

Un des rôles de l'équipe était la collecte, la mesure et le stockage des données de l'écosystème des clients media.

Mes tâches effectuées :

- **développements de Jobs Spark pour générer les metrics des parcours des sites web vers HDFS**
- développement d'une API hautement responsive (<50ms) avec un contrat Thrift
- **Intégration et configuration de kafka au sein du socle technique**
- migrations sql
- développement de jobs Java schedulés
- développement d'une application complète de génération de token
- refactoring du système des ACLs
- documentation

Compétences: Spring, spring boot2, REST, Postman, Jooq, Mockito, **Spark, Hdfs, Kafka, Docker**, Redis, Thrift, Storm, ssl two ways, GIT, MAVEN, Jenkins, Linux, Java 8, IntelliJ

Enedis

2017-2018

Devops / Data engineering

Enedis a mis en place le plus grand réseau de distribution d'électricité d'Europe. Les data scientists d'Enedis ont besoin de données structurées pour leurs analyses. **Je suis intervenu en tant que DevOps avec une casquette Data Engineer pour la mise en place d'un ETL pour intégrer, enrichir et archiver les flux de données au sein d'un Datalake basé sur Kafka et HDFS.**

Mes responsabilités :

- Scripting bash
- Customization de la configuration des brokers et connectors source Kafka (données xml, csv)
- POC d'ingestion des fichiers sources vers les brokers
- Création de jobs Azkaban pour lancer une mise à jour des données entre un NAS et une autre machine
- Mise en place de pipelines Jenkins afin de démarrer les process Kafka connect, les process de mise à jour des schémas avro
- **Pipeline Jenkins pour builder et déployer sur Openshift**
- Résultats après l'automatisation d'un accostage d'un flux de données :
2,5 Jours => 0,5 Jour
- **Accostage d'une vingtaine de flux**
- Documentation
- Développement API REST
- SSL two ways Authentification
- POC de migration de données de HDFS vers TERADATA avec des requêtes HIVE et sqoop

Compétences : Apache Kafka, Kafka connect, Jenkins, Bash, Azkaban, Java 8, REST, Openshift, HDFS, scala, hive

**Grands comptes
assurance, finance,
éditeurs**

Senior Java Développeur / Tech Lead / DevOps/ Java/REST/SGBDR/Jenkins/Gitlab-ci/Kafka/Docker/AWS

Au sein de projets grands comptes comme la Sogecap, la Coface ou d'éditeur comme

Quanthouse ou Bisam je travaille en tant que prestataire dans les domaines de l'assurance et de la finance :

- Assurance emprunteurs
- Gestion des risques d'investissements
- Api de trading
- Outils d'analyse de performance de portfolios

Mes responsabilités :

- Maintenance et les évolutions fonctionnelles des applications
- Application de la méthodologie craftsmanship pour l'amélioration de la qualité du code (sonarqube, mob programming, pair programming, tests unitaires, tests d'intégration)
- Mise en place d'un ETL basé sur Kafka Connect et Apache Kafka
- Migrations de bases de données
- Optimisation de batchs SQL
- Mise en place d'API REST
- Gestion de la CI/CD
- Mise en place de nouveaux projets
- Mise en place de microservices
- BBL

Compétences : Java 1.8 à 21, JUnit, Git, SQL, Kafka, Kafka Connect, Avro, Docker, Kubernetes, Terraform, Ansible, Jenkins, Gitlab-ci, Maven, Gradle, Spring / Spring Boot 2 / 3, SQL, ORM, Liquibase, Flyway, Oracle, PostgreSQL, Mysql, SqlServer, Tomcat, Wildfly, Weblogic, Swagger, REST, GraphQL, AWS (IAM, VPC, EC2, S3, DynamoDB, Lambda Function, Step Function, Event Bridge, appSync, Parameter Store, AWS Glue, Amazon Athena)

AWS portfolio

Exemple d'architecture :

- **Traitement "Realtime"**

Kafka producer (python) / Kafka Brokers (EC2) / Kafka Consumer / AWS S3
/ AWS Glue Crawler / Amazon Athena (SQL)

- **Batch processing :**

Airflow DAG (EC2 instance) / Python API consumer / S3

- **Développement d'une machine à état**

Environnement technique : AWS Step Function, Lambda Function, Third party API, Parameter Store, DynamoDb

- **Génération automatique d'un dashboard depuis des fichiers Csv**

[Citibike-data-analysis](#)

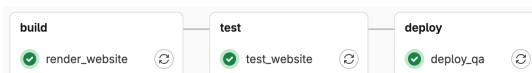
Environnement technique : docker, Python, Sql, Airflow, Metabase)

- **Développement d'un dashboard qui affiche les températures à Paris sur une plage de dates.**

[DashBoard température Paris](#)

Environnement technique : AWS IAM, Python, API, sql, EventBridge, Lambda, AWS Firehose, AWS S3, AWS Glue Crawler, AWS Workflow, AWS Glue jobs, AWS Athena, Parquet, Grafana

- Pipelines Gitlab-ci avec étapes



RÉFÉRENCES

Informations des contacts disponibles sur demande.

Jaouad B. (Indépendant) : Cloud Data engineer / DevOps

Alexandre L. (Groupe Covéa) : Tech Lead

Frédéric A. (Reacteev) : Coach agile

CERTIFICATIONS

AWS Certified Solutions Architects - Associates (SAA-CO2) chez A Cloud Guru

<https://verify.acloud.guru/FFAB157294F5>

Cloud - DevOps

<https://blent.ai/formation/cloud-devops>

Compétences : AWS Lambda · Amazon Web Services (AWS) · Terraform · Ansible · gitlab-ci · Amazon S3 · kubernetes · docker

https://www.linkedin.com/in/emmanuel-harel-dev/overlay/1635554604896/single-media-viewer?type=DOCUMENT&profileId=ACoAAACKjYEBLIZnDyLncK_laoi8tr-TrzQelo&lipi=urn%3Ali%3Apage%3Ad_flagship3_profile_view_base%3BGIG1yyf4RdqWrUiwsJ30lw%3D%3D

Data Engineer with Python

Compétences : langage python, les structures, numpy, pandas, Dataframe, ETL, airflow

<https://www.datacamp.com/completed/statement-of-accomplishment/track/46054a31a3441e51cccf8a59109e6d2f23c3b6d2>