

(+33)0751054224

Sihi.yassine@gmail.com

146 RUE DU GENERAL DE GAULLE
95120 ERMONT

07/12/1991

Yassine SIHI

BIG DATA DEVOPS



DIPLÔMES

2018

MASTER BIG DATA ET FOUILLE
DE DONNEES
Paris 8
Saint Denis

2014

LICENCE, SYSTEMES, BASES
DE DONNEES ET RESEAUX
Université Hassan 2
Casablanca

LANGUES

- Anglais
- Français
- Arabe

CERTIFICATIONS

2020

CERTIFICATION KAFKA

Confluent

- KAFKA DEV
- KAFKA ADMIN

COMPÉTENCES

Big Data	CDP, Cloudera Manager, Framework Hadoop, HDFS, Kafka, Hive, Spark, Yarn, Nifi, Oozie, Hbase, Hue, Airflow
Langages	Python, SQL, Shell
DevOps	Ansible, AWX (Ansible Tower), Git, GitHub, GitLab, K8S, Docker, AWS, Azure, IBMCloud, Terraform, Jenkins, Gitlab-ci
Outils	Splunk, Confluence, Jira, Mattermost, Vault, Teams
Méthodes	Agile Scrum
Systèmes	Windows, Mac OS, Linux Red HAT
Compétences fonctionnelles	Gestion d'incident, Conseils techniques aux équipes Incident, Support interne aux équipes incident, Astreintes

EXPÉRIENCES

SOCIETE GENERALE RESG (DATA Integrator / DEVOPS)

Février 2021 – JUIN 2024

Descriptif projet :

En tant que Data Engineer/DevOps dans un contexte anglophone, je me suis spécialisé dans la migration de la plateforme HDP vers CDP pour la Société Générale à Val de Fontenay. Ma palette de compétences englobe Cloudera Data Platform 7.1.7 avec de l'expertise dans HDFS, Yarn, Kafka, Hive, Cloudera Manager, Ranger, Oozie, Nifi, ZooKeeper, Spark, et Airflow.

Migration HDP vers CDP

- Optimisation des opérations grâce à l'automatisation via Ansible/AWX,
- Collaboration efficace avec le socle Git/GitHub
- Migration des outils de dev sur des pods K8S
- Maîtrise des développements en Python
- Développement des scripts Redhat/bash pour la migration

Montée de version des outils internes SG :

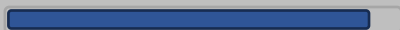
- WHATS (Annuaire de gestion des rôles et accès des projets)
- AWX (Mise à jour des jobs AWX et des rôles Ansible)
- Maîtrise de JIRA, VAULT, Confluence, GitHub pages,

Support :

- Garantir le bon fonctionnement de la plateforme Big Data
- Assurer le bon fonctionnement de applications projets
- Gestion d'incidents et de changements
- Astreintes
- Gestion des incidents complexes

EXPETISE

GIT



ANSIBLE



PYTHON



JAVA



REDHAT/BASH



CDP



KAFKA



AWS/AZURE



K8S



NIFI



RESEAUX



LOISIRS



Référent KAFKA :

- Développement des rôles Ansible pour :
 - Automatisation de la création de certificats SSL pour l'authentification auprès de Kafka
 - Automatisation du patch des topics Kafka pour les contrôles d'accès
 - Automatisation des tests de production/consommation sur des topics
- Gestion d'incidents et participation au support N1/N2
- Conseils et recommandations aux équipes projets
- Formation et workshop pour les intégrateurs BIGDATA à Paris et Bangalore
- Documentation et reporting des différents processus d'utilisation de Kafka
- Accompagnement de projets pour la consommation et production de messages depuis et vers les Process groupes NIFI

Environnement de travail :

- Langages : Python, Shell, SQL, Bash
- Systèmes : Linux Red HAT, Mac OS, Windows, Kafka
- Technologie : Ansible, GIT, K8S, Docker, API REST
- Outils : AWX (Ansible Tower), Kubernetes, VScode, GitHub, Linux, Bash, Vault, Jira, Stack CDP 7.1 (Cloudera Manager, Hive3, Spark 2, Oozie...), CDH (Kafka, NiFi), Airflow, Confluence
- Cloud : Azure, AWS

BNP PARIBAS BP2I (BIG DATA DEVOPS)

Novembre 2018 - Septembre 2020

Descriptif projet :

Au sein de l'équipe devops sur un projet KAFKA confluent, j'avais en charge la bonne gestion des clusters KAFKA.

- Contribuer à l'industrialisation du déploiement et de la maintenance de l'offre Kafka Confluent sur des VMs
- Automatisation de la création et la gestion des Topics/ACLs/Quotas/Schemas avec Ansible
- Résolution des incidents/traitements des demandes clients sur les Environnement DEV/HOMOL/PROD
- Mise à jour des composants des clusters et réalisation des tests de charge/résilience et des benches
- Créations des Dashboard Splunk pour la supervision des Clusters Kafka et intégration de flux JMX vers les métiers
- Contribuer à l'amélioration de la chaîne CI/CD pour le déploiement et la mise à jour des clusters Kafka Confluent.
- Conseil/Recommandation/Expertise des paramétrages utilisés pour optimiser les performances des clusters Kafka
- Rédaction des PV de recette et des documents sur Confluence et Wiki Teams, et réalisation des rapports journaliers de l'état des clusters de prod
- Participer à l'amélioration de l'offre Kafka par l'intégration de nouveaux éléments comme Kafka Connect et Kafka Stream.
- Contribuer à l'industrialisation du déploiement de clusters Kafka Event Streams sur le cloud IBM
- Gestion d'incidents et participation au support N2/N3 du Kafka Confluent et des équipes de développement et de data science
- Astreintes pour assurer le fonctionnement permanent de la Production

Environnement de travail :

- Langages : Python, SQL, Shell
- Kafka : Kafka confluent, Schema Registry, Rest Proxy, Kafka Connect, Kafka , Kafka stream, kafka control center
- Cloud : Event Streams, Cloud IBM
- Outils : Splunk, Git, Gitlab, Ansible, Confluence, Jira, Mattermost, Jenkins
- Autres: Apache Kafka, Gitlab, Redhat, Teams

Descriptif projet :

Au sein de l'équipe BIG DATA, j'avais en mission l'installation et la migration de clusters ZERO BIG DATA Hortonworks 2.5 vers 2.6 et la Kerbisation du cluster dans un Domain active directory.

- Approvisionnement des VMs qui vont accueillir les services HDP et Installation de HDP 2.5 avec AMBARI
- Installation/Mise à jour des packages R et redhat sur les VM à la destination des Dev.
- Développement des scripts Python et Shell de purge périodique des données sur HDFS.
- Résolution des incidents/ traitements des demandes clients

Environnement de travail :

- Langages : Python, HSQL, Shell
- Bigdata: Stuck HDP 2.5 (HDFS, Ambari, HIVE, Zookeeper, KAFKA, Ranger, Zeppelin, Views Ambari...)