

# Société Générale - Partie Administration de NIFI et Kafka

- **Gestion des mises à jour :**
  - Upgrader les versions de NIFI et Kafka.
- **Sécurité :**
  - Mise en place du système de connexion de NIFI via LDAP.
  - Gestion des ACLs des topics Kafka avec Apache Ranger.
- **Gestion des performances et de la résilience :**
  - Scale-out de Kafka en ajoutant de nouveaux brokers.
  - Utilisation de Kafka Reassign Partitions pour rééquilibrer les partitions entre brokers.
  - Augmenter le nombre de partition d'un topic Kafka pour améliorer les performances de lecture et d'écriture.
  - Configuration des réplicas des partitions de Kafka pour garantir la tolérance aux pannes.
  - Mise en place d'une politique de rétention des données dans Kafka pour éviter les problèmes d'espace disque.
- **Surveillance des métriques :**
  - Configuration de PrometheusReporting pour collecter et surveiller les métriques du cluster NIFI (par exemple, les files d'attente et le nombre de threads).
  - Mise en place de Telegraf pour collecter les métriques des machines Kafka.
  - Installation de Jolokia pour collecter les métriques des brokers Kafka.
  - Déploiement de Kafka Lag Exporter pour collecter les retards des consommateurs de Kafka pour éviter la perte de données.
  - Création de dashboards Grafana pour le monitoring complet du cluster Kafka.
- **Support des utilisateurs et optimisation des performances des applications :**
  - Accompagnement des utilisateurs dans le **tuning** des paramètres des applications qui produisent ou consomment des données dans **Kafka**.
  - Surveillance des **logs Kafka** pour détecter les erreurs critiques.

## **SNCF - Partie Administration et monitoring du cluster HDF (Hortonworks Data Flow) composé de NIFI, Kafka, Zookeeper ...**

- **Installation et sécurisation** : Installation de NIFI et Kafka via Ambari Server (inclus dans HDF), sécurisation de NIFI et Kafka en mode SSL et SASL\_SSL respectivement avec keytool.
- **Gestion des flux NIFI** : Installation de NIFI Registry pour la gestion du versioning des flux NIFI.
- **Support du cluster NIFI** : Gestion du support du cluster NIFI, incluant la gestion de différents aspects comme la cohérence du fichier flow.xml en cas de déconnexion d'un nœud, ainsi que d'autres processus critiques pour maintenir la disponibilité du cluster.
- **Tuning de NIFI** : Optimisation des paramètres de NIFI pour améliorer les performances, tels que l'ajustement de la mémoire JVM, le réglage des tailles de buffer, et la gestion des threads et des ressources système pour garantir une efficacité maximale.
- **Monitoring et reporting** : Mise en place AmbariReportingTask pour le monitoring du cluster NIFI.

## **EDF - Partie Administration de NIFI :**

- **Mises à jour et gestion des versions** :
  - Étudier les impacts des montées de version de NIFI.
  - Mettre à jour les versions de NIFI
- **Déploiement et personnalisation des processeurs** :
  - Déployer des custom processors.
- **Intégration et synchronisation** :
  - Synchroniser NIFI et GitHub avec RegistryClientGithub.
- **Surveillance et métriques** :
  - Mettre en place PrometheusReporting pour surveiller le cluster NIFI.
- **Optimisation des performances** :

- Effectuer le tuning des paramètres de NIFI.