

Proposition d'Offre - Accélération Mise en Place Dataplateforme (1 Semaine)

Client : CBAO

Projet : Intervention intensive d'une semaine pour accélérer la mise en place de la data plateforme

Date : 12/12/2025

Consultant : Bassirou KA

Sociétés : Tecsen France / Back Consulting Sénégal

Email : ka.bassirou@gmail.com

Téléphone : +33 6 34 24 79 25

Sommaire

1. [Contexte et objectifs](#)
 2. [Architecture de référence](#)
 3. [Détail des prestations](#)
 4. [Planification](#)
 5. [Chiffrage](#)
 6. [Modalités d'intervention](#)
 7. [Livrables](#)
 8. [Profil du consultant](#)
 9. [Conditions générales](#)
-

1. Contexte et objectifs

Contexte

La CBAO souhaite accélérer la mise en place de sa dataplateforme moderne sur infrastructure Kubernetes on-premise. Cette intervention intensive d'une semaine se concentre sur les composants essentiels pour l'intégration et le traitement de données.

Objectifs

Cette intervention permettra de :

- **Valider et optimiser** l'architecture proposée de la data plateforme
 - **Déployer et configurer** Apache Nifi en production (cluster HA) pour l'intégration de données
 - **Déployer et configurer** Nifi Registry pour la gestion de versions des flows
 - **Mettre en place** l'environnement Spark sur Kubernetes avec outils et templates de déploiement
 - **Transférer les compétences** à l'équipe Data Engineer et DevOps
 - **Documenter** les procédures de déploiement pour Apache Airflow et Trino
-

2. Architecture de référence

Composants ciblés

Intégration de données :

- **Apache Nifi** : Intégration de données en temps réel et batch
- **Apache Nifi Registry** : Gestion de versions des flows Nifi

Traitement et transformation :

- **Apache Spark** : Traitement distribué de grandes volumétries sur Kubernetes
- **Spark History Server** : Visualisation et monitoring des jobs Spark

Orchestration (documentation) :

- **Apache Airflow** : Orchestration et planification des workflows de données
 - **Trino** : Moteur SQL universel pour accès fédéré aux données
-

3. Détail des prestations

Phase 1 : Intervention présente à Dakar (5 jours)

Objectif : Déployer et configurer les composants essentiels avec transfert de compétences

Modalité : 100% présentiel à Dakar

Détail des activités :

Activités obligatoires (4 jours)

1.1 - Revue et Challenge de l'architecture (0.5 jour)

- Analyse de l'architecture de la data plateforme
- Validation et optimisation des choix technologiques
- Recommandations architecturales
- Plan de déploiement détaillé

1.2 - Mise en place d'Apache Nifi et Nifi Registry (2.5 jours)

- Déploiement d'Apache Nifi en cluster HA sur Kubernetes
- Configuration de ZooKeeper pour la coordination du cluster
- Déploiement et configuration de Nifi Registry
- Intégration Nifi Registry avec Nifi pour la gestion de versions
- Configuration de la haute disponibilité
- Tests de validation et démonstration
- Transfert de compétences à l'équipe

1.3 - Installation de l'environnement Spark sur Kubernetes (1 jour)

- Déploiement du Spark History Server
- Mise en place d'outils et templates de déploiement d'applications Spark clusterisées sur Kubernetes
- Configuration des ressources et des paramètres Spark
- Démonstration sur une application Spark

- Documentation des procédures de déploiement
- Transfert de compétences à l'équipe

★ Activités "Nice to have" (1 jour - selon temps disponible)

1.4 - Documentation des procédures de déploiement (1 jour)

- Documentation d'une procédure de mise en place d'Apache Airflow
- Documentation d'une procédure de mise en place de Trino
- Revue des procédures avec l'équipe DevOps
- Validation et ajustements

Livrables Phase 1 :

- Manifests Kubernetes / Helm charts pour Nifi, Nifi Registry et Spark
- Documentation de configuration complète
- Scripts d'installation et de déploiement
- Templates de déploiement d'applications Spark
- Documentation des procédures Airflow et Trino (si temps disponible)
- Application Spark de démonstration

Phase 2 : Intervention support à distance (5 jours)

Objectif : Finaliser et optimiser les composants déployés

Modalité : 100% distanciel

Détail des activités :

2.1 - Configuration avancée de Nifi (1 jour)

- Optimisation des configurations Nifi
- Configuration avancée des processors
- Paramétrage de la performance et de la scalabilité
- Documentation des bonnes pratiques

2.2 - Installation d'Apache Airflow (2 jours)

- Déploiement d'Apache Airflow sur Kubernetes

- Configuration de la base de données (PostgreSQL)
- Configuration de l'authentification (Keycloak si disponible)
- Tests et validation
- Documentation complète

2.3 - Installation du cluster Trino (2 jours)

- Déploiement du cluster Trino sur Kubernetes
- Configuration des connecteurs de données
- Configuration de la sécurité et de l'authentification
- Tests et validation
- Documentation complète

Livrables Phase 2 :

- Manifests Kubernetes / Helm charts pour Airflow et Trino
 - Documentation de configuration complète
 - Scripts d'installation et de déploiement
 - Guide d'optimisation et de bonnes pratiques
 - Documentation d'intégration avec les autres composants
-

4. Planification

Durée totale : 2 semaines (10 jours ouvrés)

Période	Activités principales
Semaine 1	Phase 1 - Intervention présente à Dakar (5 jours)
Semaine 2	Phase 2 - Support à distance (5 jours)

Jalons :

- **J1** (Fin Semaine 1) : Nifi, Nifi Registry et Spark opérationnels
 - **J2** (Fin Semaine 2) : Airflow et Trino déployés et opérationnels
 - **Semaine 1** : 5 jours ouvrés en présentiel à Dakar
 - Déploiement des composants essentiels
 - Formation et transfert de compétences
 - Tests de validation
 - **Semaine 2** : Travail à distance
 - Support et déploiement des composants complémentaires
 - Documentation et optimisation
-

5. Chiffrage

Répartition des jours-hommes

Phase	Prestation	Durée (JH)
Phase 1.1	Revue et Challenge de l'architecture	0.5
Phase 1.2	Apache Nifi et Nifi Registry	2.5
Phase 1.3	Environnement Spark sur Kubernetes	1
Phase 1.4	Documentation Airflow et Trino (Nice to have)	1
Phase 2.1	Configuration avancée Nifi	1
Phase 2.2	Installation Apache Airflow	2
Phase 2.3	Installation cluster Trino	2
TOTAL		10 JH

Détail du chiffrage

Répartition par phase :

- Phase 1 (présentiel - 5 jours) : 1 750 000 FCFA HT (5 JH × 350 000 FCFA/JH)
- Phase 2 (distanciel - 5 jours) : 1 750 000 FCFA HT (5 JH × 350 000 FCFA/JH)

Récapitulatif financier

Poste	Montant HT
Phase 1 - Intervention présentelle (5 JH)	1 750 000 FCFA
Phase 2 - Support à distance (5 JH)	1 750 000 FCFA
TOTAL HT	3 500 000 FCFA
TVA (20%)	700 000 FCFA
TOTAL TTC	4 200 000 FCFA

Tarif unitaire : 350 000 FCFA / jour-homme (applicable aux deux phases)

Conditions de facturation

- **Phase 1** : Facturation à la fin de l'intervention présentelle (1 750 000 FCFA HT)
 - **Phase 2** : Facturation à la livraison finale (1 750 000 FCFA HT)
-

6. Modalités d'intervention

Phase	Présentiel	Distanciel	Total
Phase 1	5 JH	0 JH	5 JH
Phase 2	0 JH	5 JH	5 JH
TOTAL	5 JH	5 JH	10 JH

Présentiel (Semaine 1) : Déploiements critiques, formation et transfert de compétences, tests de validation

Distanciel (Semaine 2) : Déploiement des composants complémentaires, documentation, support et optimisation

Communication : Points d'avancement quotidiens (présentiel) puis selon besoin (distanciel) via Teams, Git, Confluence/Wiki

7. Livrables

Phase 1 :

- Manifests Kubernetes / Helm charts (Nifi, Nifi Registry, Spark)
- Documentation de configuration et scripts d'installation
- Templates de déploiement Spark
- Application Spark de démonstration
- Documentation Airflow et Trino (si temps disponible)

Phase 2 :

- Manifests Kubernetes / Helm charts (Airflow, Trino)
- Documentation de configuration complète
- Scripts d'installation et guides d'optimisation
- Documentation d'intégration

Transversaux : Documentation technique complète, Architecture as Code, procédures de maintenance

8. Profil du consultant

Consultant : Bassirou KA

Sociétés : Tecsen France / Back Consulting Sénégal

Architecte Data Ops avec 14 ans d'expérience en architecture de données et dataplateformes.

Expérience : 14 ans en architecture de données et data engineering, expertise en dataplateformes cloud-native, spécialisation Kubernetes on-premise et cloud, projets grands comptes en France et Afrique

Compétences : Apache Airflow, Spark, Nifi, DBT, Trino, Kubernetes, Helm, Terraform, Keycloak, OPA, Prometheus, Grafana

Certifications : CKA (Certified Kubernetes Administrator), ISO-27001, AWS/Azure

9. Conditions générales

Prérequis

Infrastructure :

- Cluster Kubernetes fonctionnel et opérationnel
- Accès en tant qu'administrateur au cluster
- Provider de stockage disponible (OpenEBS ou équivalent)
- Point d'entrée des services configurés (MetalLB, external load balancer ou équivalent)
- Identity Provider disponible (Keycloak ou autre) pour l'authentification
- Accès internet pour téléchargement d'images Docker

Équipe :

- Équipe Data Engineer disponible pour le transfert de compétences
- Équipe DevOps disponible pour la revue des procédures
- Accès aux environnements de développement/staging/production
- Accès aux sources de données pour tests (si nécessaire)

Outils :

- Accès Git pour versioning des configurations
- Outils de communication (Teams/Slack)
- Accès à la documentation existante
- Accès VPN (pour la phase distancielle)

Gestion des risques

Risques identifiés :

1. Complexité de l'infrastructure Kubernetes

- *Mitigation* : Expérience confirmée, documentation détaillée, tests préalables

2. Compatibilité des composants

- *Mitigation* : Versions validées, tests en environnement de développement

3. Disponibilité de l'équipe

- *Mitigation* : Planning flexible, communication régulière

4. Temps disponible pour les activités "Nice to have"

- *Mitigation* : Priorisation des activités obligatoires, activités complémentaires selon disponibilité

Propriété intellectuelle

- Les livrables (code, configurations, documentation) sont la propriété de la CBAO
- Le consultant conserve le droit d'utiliser les connaissances acquises (sans divulguer d'informations confidentielles)

Confidentialité

- Engagement de confidentialité sur les informations de la CBAO
 - Respect de la réglementation applicable sur les données
-