

# Dhia Eddine Gharsallaoui

Paris, Île-de-France, France | gharsallaoui.dhia@gmail.com | 07 58 77 03 14 | [LinkedIn](#) | [Portfolio](#)

## RÉSUMÉ

Distributed Platform Architect avec 3 ans d'expérience, spécialisé dans la conception et le déploiement de plateformes de données distribuées pour diverses entreprises. Expert en architecture de données cloud-native, optimisation des performances et automatisation. Maîtrise approfondie des technologies de traitement de données massives incluant Apache Spark, Delta Lake, et Trino. Expérience significative dans l'implémentation de solutions data mesh et la création de plateformes de données unifiées multi-sources. Contributeur actif aux projets open-source dans le domaine du data engineering.

## EXPÉRIENCE PROFESSIONNELLE

### Architecte Solutions

*Skilld | Février 2022 - Présent | Paris, France*

- Conception et déploiement de plateformes de données distribuées pour multiples clients, implémentant des architectures data mesh et des pipelines de traitement de données évolutifs.
- Configuration, déploiement et maintenance de clusters Kubernetes haute disponibilité multi-environnements (dev, staging, prod) avec mise en place de politiques de sécurité, gestion des backups et plans de reprise d'activité.
- Déploiement et maintenance de clusters Kafka distribués, monitoring des performances, et optimisation des configurations pour gérer des charges importantes de streaming de données.
- Mise en place de Trino comme solution de requêtage unifié, permettant aux clients d'interroger efficacement leurs données à travers multiples sources (Delta Lake, Apache Druid, PostgreSQL, Prometheus) avec des performances optimisées.
- Développement et optimisation de pipelines de traitement de données utilisant Apache Spark et Delta Lake, implémentant des contrôles de qualité des données et une gouvernance robuste.
- Implémentation d'une architecture de monitoring complète avec Prometheus et Grafana, incluant des tableaux de bord personnalisés pour le suivi des performances des plateformes de données.
- Conception et maintenance d'une infrastructure Kubernetes multi-cluster pour le déploiement automatisé des plateformes de données.
- Automatisation des processus de déploiement et de configuration utilisant Ansible et GitLab CI, assurant une cohérence et une reproductibilité des environnements.
- Implémentation d'un système de traçage distribué avec OpenTelemetry pour une visibilité complète des pipelines de données.
- Implémentation d'une couche de sécurité complète avec intégration SSO utilisant l'écosystème Ory, gérant l'authentification et l'autorisation à travers plusieurs applications.

### Développeur Blockchain Freelance

*ChainsAtlas | Mars 2024 - Présent*

- Développement et optimisation de protocoles de bridge multi-chain intégrant les chaînes compatibles EVM avec XRPL, implémentant des transferts d'actifs cross-chain sécurisés et des opérations NFT.
- Architecture et implémentation de systèmes de récompenses de bataille avec validation multi-signature pour le gaming blockchain, incorporant des mécanismes de distribution de récompenses XRP et NFT.
- Construction de systèmes de traitement de transactions robustes avec logique de retry et gestion des nonces, atteignant une meilleure fiabilité des transactions à travers plusieurs réseaux blockchain.
- Conception et implémentation d'opérations blockchain agnostiques impliquant des smart contracts en Solidity et TypeScript, avec un focus sur les fonctionnalités de wallet multi-signature sécurisées.
- Amélioration de l'infrastructure de bridging blockchain en développant des systèmes de vérification et de gardiennage, implémentant des protocoles de validation sécurisés pour les opérations cross-chain.

### Ingénieur Logiciel Freelance

*idatase GmbH | Janvier 2024 - Mars 2024 | Télétravail*

- Développement de pipeline d'ingestion de données évolutif utilisant Python/Flask et PostgreSQL.
- Implémentation d'architecture microservices containerisée utilisant Docker Compose et Kubernetes.
- Création de pipeline de test et déploiement automatisé utilisant GitHub Actions.
- Conception et implémentation d'un système de collecte de métriques en temps réel.

## COMPÉTENCES

### Plateformes de Données

Trino, Apache Spark, Delta Lake, Apache Druid, Data Mesh, Apache Atlas, dbt, Apache Airflow

### Infrastructure & Automatisation

Kubernetes, Docker, Ansible, GitHub Actions, GitLab CI, Helm, Terraform, Infrastructure as Code

### Monitoring & Observabilité

Prometheus, Grafana, Elasticsearch, Graylog, Jaeger, OpenTelemetry, AlertManager

### Bases de Données & Cache

PostgreSQL, TimescaleDB, Redis, DragonflyDB, Trino, Apache Cassandra

### Programmation & Blockchain

Go (Golang), Python, Scala, Solidity, TypeScript, Smart Contracts, Protocoles Cross-chain

### Sécurité & Conformité

Cert-Manager, RBAC, OAuth2/OIDC, Ansible Vault, Écosystème Ory

## PROJETS OPEN SOURCE

### Mage-ai - [Open Source](#)

Implémentation de l'authentification LDAP, intégration Apache Druid et système de logging centralisé.

### Go-mail - [Open Source](#)

Développement du pattern d'architecture middleware et implémentation du middleware de chiffrement OpenPGP avec couverture de tests complète.

### Go-RTE - [Open Source](#)

Création d'une bibliothèque client Go pour les APIs RTE avec une façon propre et uniforme d'interagir avec différents endpoints API.

## FORMATION

### Master en Informatique

*MINES ParisTech | 2020 - 2021*

*Mineure en Mathématiques | Spécialisation en Systèmes Distribués et Cloud Computing*

### Diplôme d'Ingénieur en Génie Industriel

*ENIT | 2018 - 2021*

*Mineure en Data Science | Spécialisation en Big Data Analytics*