

## NOTRE HISTOIRE :

🇪🇺 Rejoignez Scaleway et façonnez le cloud souverain de demain !

Depuis 1999, nous concevons des infrastructures sécurisées, durables et pensées pour accompagner les entreprises les plus ambitieuses.

Historiquement reconnus pour nos serveurs dédiés Dedibox, nous avons pris en 2015 un virage stratégique vers le cloud computing. En restant fidèles à nos principes de simplicité, flexibilité et d'excellence technique, nous avons su trouver notre place parmi les leaders européens du secteur.

Avec l'essor de l'intelligence artificielle, nous avons choisi de renforcer notre engagement en bénéficiant du soutien du groupe Iliad, qui investit aujourd'hui 3 milliards d'euros pour développer une alternative IA souveraine et massive face aux géants américains et asiatiques.

Chaque jour et grâce à notre riche catalogue de produits et services (bare metal, conteneurisation, serverless, IA, etc.), Scaleway sert fièrement 38 000 clients privés et publics, de Potoroom à Mistral AI en passant par Golem.AI ou l'ADEME.

---

## POURQUOI AVONS-NOUS BESOIN DE VOUS ?

Notre croissance nous pousse à renforcer notre équipe **AI Infrastructure** afin de soutenir le développement de notre plateforme d'inférence de nouvelle génération.

Votre mission sera de **concevoir et d'opérer une infrastructure d'inférence de production pour les workloads IA**, permettant le déploiement et la scalabilité des LLMs et des applications d'IA générative sur un cloud européen souverain.

---

## VOTRE FUTUR ÉQUIPE :

Nous travaillons dans un environnement collaboratif et international, où la diversité des Scalars combinée à l'esprit de partage permettent de faire émerger de nouveaux projets chaque jour et d'avancer ensemble dans la concrétisation de nos ambitions.

Vous ferez partie d'une équipe de 4 DevOps engineers, travaillant sur les sujets backend et infrastructure liés à l'inférence IA. Cette squad fait partie d'une **AI Tribe** plus large, organisée autour de 4 produits stratégiques :

- **Managed Inference** : une plateforme pour déployer, scaler et monitorer des modèles IA en production
- **Generative APIs** : une couche d'API unifiée pour accéder aux modèles génératifs les plus avancés (LLMs, Diffusion, etc.)

- **AI HPC** : des clusters massifs, avec des milliers de GPU, dédiés aux clients ayant de forts besoins d'entraînement de modèles
- **AI Tooling** : développement de technologies internes pour booster Scaleway et le groupe Iliad au quotidien

 **Manager** : Vous reporterez à **Grégoire de Turckheim**, chez Scaleway depuis plus de 14 ans. Son parcours riche (ingénierie, product management, logistique, achats, partenariats stratégiques) nourrit un style de management basé sur la vision long terme, l'agilité et la culture de l'ownership.

---

## VOTRE QUOTIDIEN :

- Construire et opérer l'infrastructure permettant de servir des LLMs et autres modèles génératifs à grande échelle
  - Concevoir, déployer et maintenir des services Kubernetes optimisés pour l'inférence IA
  - Développer en **Golang** pour construire des services backend robustes et performants
  - Optimiser les stacks d'inférence avec des outils comme **vLLM, Triton et CUDA**
  - Intégrer des outils IA open source tels que **Hugging Face, KServe** ou des composants custom
  - Participer aux discussions d'architecture et prendre des décisions techniques à fort impact
  - Rester à jour sur l'évolution rapide des écosystèmes d'IA appliquée à l'inférence
  - Débuguer et résoudre des problématiques complexes d'infrastructure et de déploiement
  - Collaborer avec les équipes internes pour intégrer de nouvelles fonctionnalités IA
  - Contribuer aux revues de code et au partage de connaissances au sein de l'équipe
-

## À PROPOS DE VOUS :

### HARDSKILLS

- Solide expérience en **Golang** dans des environnements de production
- Maîtrise de **Kubernetes** et des outils d'orchestration de conteneurs
- Connaissances des modèles IA (LLMs, modèles de diffusion, ...)
- Une expérience concrète dans le déploiement et la mise à l'échelle de modèles ML est un vrai plus
- Bonne compréhension des architectures backend et des pratiques DevOps

### SOFTSKILLS

- Capacité d'apprentissage rapide et forte adaptabilité aux nouvelles technologies
- Curiosité marquée et envie d'explorer en profondeur les nouveaux outils IA
- Esprit proactif, prêt à prendre des initiatives
- Autonomie technique et ownership de votre périmètre
- Communication claire, avec une capacité à vulgariser des systèmes complexes

---

## CE QUE VOUS TROUVEREZ CHEZ SCALEWAY ++++

- **Travail hybride** : jusqu'à 3 jours de télétravail par semaine
- **Locaux modernes** : espaces créatifs et ergonomiques, bien situés, avec terrasses et stationnements vélos
- **Restauration** : repas préparés par notre chef au siège ; carte Swile pour les sites régionaux
- **Bien-être** : prise en charge pour salle de sport, garde d'enfants et services aux aidants
- **Culture internationale** : l'anglais est aussi courant que le français, avec des dizaines de nationalités représentées

- **Mobilité interne** : fortes opportunités de carrière, avec passerelles possibles vers d'autres entités du groupe Iliad
- 

### Pourquoi rejoindre l'aventure Scaleway ?

- ✓ Une offre produit riche et diversifiée : plus de 100 produits cloud public sur le IaaS, PaaS et IA
  - ✓ Un environnement technique de pointe : infrastructures modernes, serveurs bare metal haute performance, défis techniques stimulants
  - ✓ Un engagement pour un cloud responsable : datacenters alimentés uniquement par de l'énergie renouvelable depuis 2017, empreinte écologique réduite, certification au plus haut niveau
- 

### LA SUITE ...

- Call de découverte avec un recruteur (30 min)
- Entretien technique pour valider vos expertises (1h)
- Exercice pratique de **problem-solving opérationnel**
- Entretien avec le Head of de la Tribe pour approfondir vos échanges et valider votre fit avec l'équipe (45 min)
- Entretien RH pour visiter nos locaux et échanger avec vos futurs collègues

*Chez Scaleway, nous nous engageons à construire un environnement inclusif et respectueux où chacun a une juste opportunité de s'épanouir.*

*Toutes les candidatures sont étudiées avec soin, sans distinction d'âge, de genre, d'orientation sexuelle, d'origine ethnique ou sociale, de religion, de handicap ou de toute autre caractéristique.*

*Nous croyons que les grandes idées viennent de partout et de chacun — c'est pourquoi vous devriez absolument postuler.*