



Offre n°2022-05111

Ingénieur Exploration et recommandation pour la réutilisation d'images de conteneurs (H/F)

Type de contrat : CDD

Niveau de diplôme exigé : Thèse ou équivalent

Fonction : Ingénieur scientifique contractuel

Niveau d'expérience souhaité : Jeune diplômé

A propos du centre ou de la direction fonctionnelle

Le centre de recherche Inria Lille – Nord Europe, créé en 2008, compte 360 personnes dont 305 scientifiques répartis dans 15 équipes de recherche. Reconnu pour son implication forte dans le développement socio-économique sur le territoire des Hauts-De-France, le centre de recherche Inria Lille – Nord Europe poursuit une démarche de proximité avec les grandes entreprises et les PME. En favorisant ainsi les synergies entre chercheurs et industriels, Inria participe au transfert de compétences et d'expertises dans les technologies numériques et donne accès aux meilleures recherches européennes et internationales au bénéfice de l'innovation et des entreprises notamment en région.

Depuis plus de 10 ans, le centre Inria Lille – Nord Europe est installé au coeur de l'écosystème universitaire et scientifique lillois ainsi qu'au coeur de la Frenchtech avec un showroom technologique, basé avenue de Bretagne à Lille, sur le site d'excellence économique consacré aux technologies de l'information et de la communication (TIC) qu'est EuraTechnologies.

Contexte et atouts du poste

Cet(te) ingénieur(e) s'inscrit dans une collaboration entre l'équipe de recherche Spirals (<https://team.inria.fr/spirals>) et la société Berger - Levrault (<https://www.berger-levrault.com>).

L'objectif est de développer une chaîne d'outils pour la réutilisation d'images de conteneurs Docker. Cette chaîne comprendra un outil d'exploration de descripteurs Docker Compose ainsi qu'un outil de recommandation.

L'ingénieur sera principalement hébergé au sein de la direction de la recherche et innovation de Berger – Levrault à Limonest (69) et visitera l'équipe Spirals de Lille quelques jours par mois. Les frais de déplacements seront pris en charge dans la limite du barème en vigueur.

Mission confiée

Pour faciliter le déploiement logiciel au sein des infrastructures cloud, les technologies de conteneurs [1] comme Docker, Rocket, LXC, LXI, Linux-VServer, OpenVZ, runC, FreeBSD Jail, Oracle Solaris Zone, Podman, Kubernetes, pour n'en citer que quelques-unes, s'imposent progressivement. Ces dernières assistent le conditionnement, la livraison et le déploiement d'applications métiers et s'inscrivent dans la démarche DevOps. Les conteneurs sont aussi un des piliers techniques pour supporter des architectures micro-services permettant de migrer les applications monolithiques vers des applications réparties dans les nuages passant mieux à l'échelle et modulairement maintenables.

Actuellement, Docker s'est imposé comme la technologie leader du domaine. Docker fournit un format de conditionnement des images (binaires) de conteneurs, le dépôt Docker Hub pour la livraison d'images de conteneurs et le langage Docker Compose pour la description et le déploiement d'architectures micro-services. A titre illustratif au 31 mai 2022, au moins 9 135 203 images sont librement disponibles sur Docker Hub [2] et au moins 1 704 125 descripteurs Docker Compose sont librement disponibles sur GitHub [3].

Pour déployer une image Docker, il est toujours nécessaire de configurer le conteneur qui hébergera cette image via par exemple des variables d'environnement, des dépendances avec d'autres conteneurs et des paramètres d'exécution du conteneur comme les ressources CPU, mémoire, IO à utiliser, les capacités systèmes, la configuration réseau, les volumes à monter, etc. L'ensemble de ces éléments de configuration est exprimable avec le langage Docker Compose [4].

Le problème est que cette configuration n'est jamais spécifiée formellement. Elle est quelquefois informellement documentée. Mais la plupart du temps, un développeur s'inspire d'exemples disponibles sur le réseau, sur GitHub ou autres. Mais alors le développeur ne connaît jamais l'ensemble exhaustif des éléments de configuration d'une image de conteneur, il peut s'inspirer d'exemples erronés ou défectueux

reproduisant ainsi ces erreurs et défauts et enfin il ne peut garantir aucune qualité sur la réutilisation des images.

La solution envisagée consistera à explorer (data mining) automatiquement des descripteurs Docker Compose afin de construire un modèle probabiliste des éléments de configuration associés à chaque image Docker. Pour cela, nous pourrions au minimum explorer les 1 704 125 fichiers Docker Compose disponibles sur GitHub [3]. Le modèle probabiliste à construire capturera pour chaque image Docker la probabilité de devoir configurer une certaine caractéristique, par ex. variables d'environnement, ressources et capacités système, dépendances vers d'autres conteneurs, réseaux et volumes. Le fait d'alimenter ce modèle via un très grand nombre de descripteurs devrait statistiquement réduire l'impact des données erronées ou défectueuses. Ce modèle pourra ainsi être vu comme une formalisation de la configuration des images Docker construite par apprentissage automatique. Enfin, ce modèle permettra d'élaborer des recommandations pour améliorer les descripteurs Docker Compose produits par des développeurs de la société Berger Levrault (ou d'ailleurs).

[1] <https://www.ionos.fr/digitalguide/serveur/know-how/apercu-des-alternatives-a-docker/>

[2] <https://hub.docker.com/search?q=&type=image>

[3] <https://github.com/search?q=extension%3Ayaml+extension%3Ayml+filename%3Adocker-compose&type=code>

[4] <https://docs.docker.com/compose/compose-file/>

Principales activités

L'ingénieur devra dans un premier temps élaborer un état de l'art des approches de data mining et d'apprentissage automatique. Dans un second temps, le post-doctorant construira le modèle probabiliste sous-jacent à l'approche proposée puis il développera le logiciel d'exploration automatique de descripteurs Docker Compose afin d'alimenter le modèle ainsi que l'outil de recommandation exploitant le modèle alimenté. Enfin, l'ingénieur publiera ce travail dans l'une des meilleures revues du domaine. Les logiciels de Berger-Levrault serviront de cas d'évaluation à une telle technologie de recommandation de descripteurs d'infrastructure. Le moment est particulièrement opportun car la société Berger-Levrault s'inscrit dans une dynamique de modernisation de ses logiciels pour les conduire vers des environnement cloud.

Compétences

Compétences techniques et niveau requis : Maîtrise en data mining, analyse statistique et développement logiciel

Langues : Français et anglais

Avantages

- Vous intégrerez une équipe dynamique composée d'experts scientifiques internationaux dans le domaine des systèmes répartis et du génie logiciel (<https://team.inria.fr/spirals/>);
- Vous travaillerez sur des activités de recherche émergents avec des acteurs internationaux reconnus du *cloud computing* dans le contexte de collaborations et projets européens de l'équipe Spirals;
- Vous travaillerez dans un environnement de travail stimulant et agréable (participation transport (50%), restauration sur site ; télétravail ; congé et autorisation spéciale absence (45 jours), équipement visio conférence, laboratoire technique d'expérimentation...);
- Vous pourrez bénéficier de formations de qualité adaptées à vos besoins et appétences, aussi bien techniques que méthodologiques ou linguistiques;
- Au-delà de l'amélioration de vos compétences techniques, Inria vous offre l'opportunité de développer vos compétences entrepreneuriales en participant à des événements de sensibilisation et à des formations sur la création de start-up (horizon start up, formation propriété intellectuelle, hackAthon... <https://www.inria.fr/fr/inria-startup-studio>);
- Pour les candidats internationaux, nos services administratifs vous accompagnent dans les différentes démarches administratives (visa, titre de séjour, sécurité sociale, logement, banque...).

Rémunération

2 653€ brut mensuel

Informations générales

- **Thème/Domaine** : Systèmes distribués et intergiciels
Système & réseaux (BAP E)
- **Ville** : Limonest (69)
- **Centre Inria** : [Centre Inria de l'Université de Lille](#)
- **Date de prise de fonction souhaitée** : 2022-10-01
- **Durée de contrat** : 2 ans
- **Date limite pour postuler** : 2022-08-31

Contacts

- Équipe Inria : [SPIRALS](#)
- Recruteur :
Merle Philippe / Philippe.Merle@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

L'essentiel pour réussir

- Avoir une expertise en data mining, analyse statistique et développement logiciel
- Aimer travailler en équipe est une qualité essentielle pour réussir cette mission
- Se sentir à l'aise dans un environnement industriel et scientifique

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

CV + lettre de motivation + lettres de recommandation

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.