



Offre n°2022-05221

Doctorant F/H Management of mutable data over P2P storage

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Doctorant

Contexte et atouts du poste

This PhD thesis will be in the context of a collaboration between HIVE and Coast and Myriads Inria teams. The Ph.D student will be located at Inria Nancy-Grand Est and will be visiting team Myriads at Inria Center of the University of Rennes and the Hive offices in Cannes.

About Hive:

Hive intends to play the role of a next generation cloud provider in the context of Web 3.0. Hive aims to exploit the unused capacity of computers to offer the general public a greener and more sovereign alternative to the existing clouds where the true power lies in the hands of the users. It relies both on distributed peer-to-peer networks, on the encryption of end-to-end data and on blockchain technology.

About Inria Nancy - Grand Est:

The Inria Nancy - Grand Est center is one of Inria's eight centers and has twenty project teams, located in Nancy, Strasbourg and Saarbrücken. Its activities occupy over 400 people, scientists and research and innovation support staff, including 45 different nationalities. The Inria Center is a major and recognized player in the field of digital sciences. It is at the heart of a rich R&D and innovation ecosystem: highly innovative PMEs, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institutes, etc.

About Inria Center of the University of Rennes:

The Inria Center of the University of Rennes is one of Inria's eight centers and has more than thirty research teams. The Inria Center is a major and recognized player in the field of digital sciences. It is at the heart of a rich R&D and innovation ecosystem: highly innovative PMEs, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institutes, etc.

Mission confiée

For availability and performance reasons, data is replicated. Several users have to be able to update concurrently the replicas of the same data without losing their modifications. Hive solution relies on IPFS (<https://ipfs.io/>) and mutable data support is offered by means of the mutable file system API of IPFS. However, there is no support for merging concurrent changes.

Data replication also raises the question about replicas placement. Depending on the nature and use of the data, placement is a critical issue. Many data-based applications performances are improved when data is local. Data locality can be approached in two ways: either considering the placement of data as input and then try to place tasks accordingly, or trying to place data to improve the future jobs based on this data. Replica placement also has its importance when considering data consistency. In highly distributed environments, maintaining consistency between data has a cost which depends on the distance between replicas.

In this Ph.D. thesis, we plan to propose a replication mechanism over sharded encrypted data that merges concurrent changes and that optimizes the cost of this merging by a suitable replica placement. We propose using CRDTs (Conflict-free Replicated Data Types) [2, 3] as replication mechanism as they are suitable for end-to-end encryption in a peer-to-peer environment where data will be decrypted only at the receiver side and conflicts can be resolved locally. There is therefore no need to decrypt data during data transmission as it is the case for centralised architectures where servers require un-encrypted data in order to perform merging. The challenge in this approach is to develop CRDTs on sharded data stored on IPFS.

References:

[1] J. Benet. "IPFS - Content Addressed, Versioned, P2P File System". In: CoRR abs/1407.3561 (2014). doi: 10.48550/arXiv.1407.3561. arXiv: 1407.3561.

[2] M. Shapiro, N. M. Preguic a, C. Baquero, and M. Zawirski. "Conflict-Free Replicated Data Types". In: 13th International Symposium on Stabilization, Safety, and Security of Distributed Systems, SSS 2011. Oct. 2011, pp. 386–400. doi: 10.1007/978-3-642-24550-3_29.

[3] L. Andre , S. Martin, G. Oster, and C.-L. Ignat. "Supporting adaptable granularity of changes for massive-scale collaborative editing". In: Proceedings of the International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom 2013). Austin, Texas, USA, Oct. 2013

Principales activit es

The two main replication mechanisms based on CRDTs, operation-based and state-based, have both advantages and disadvantages. On the one hand, operation-based synchronisation is more efficient since one only needs to ship small updates, but it requires exactly once causally-ordered broadcast. On the other hand, state-based synchronisation requires only reliable broadcast, but it induces communication overhead by shipping the whole state. The first step would be the study of these replication mechanisms and the evaluation of their suitability for the context of HIVE.

The next step would be to study the feasibility of the combination of replica placement algorithms with replication mechanisms based on CRDTs. The combined replica placement algorithms with consistency maintenance algorithms based on CRDTs has to be evaluated in terms of time and space complexity and network communication costs.

Avantages

- Restauration subventionn e
- Transports publics rembours s partiellement
- Cong s: 7 semaines de cong s annuels + 10 jours de RTT (base temps plein) + possibilit  d'autorisations d'absence exceptionnelle (ex : enfants malades, d m nagement)
- Possibilit  de t l travail (apr s 6 mois d'anciennet ) et am nagement du temps de travail
-  quipements professionnels   disposition (visioconf rence, pr ts de mat riels informatiques, etc.)
- Prestations sociales, culturelles et sportives (Association de gestion des  uvres sociales d'Inria)
- Acc s   la formation professionnelle
- S curit  sociale

R mun ration

1982,00  brut mensuel les deux premi res ann es (1594,00  net)

2085,00  brut mensuel la 3 me ann e (1677,00  net)

1982,00  gross monthly for the first two years (1594,00  net)

2085,00  gross monthly the 3rd year (1677,00  net)

Informations g n rales

- **Ville** : Villers l s Nancy
- **Centre Inria** : [Centre Inria de l'Universit  de Lorraine](#)
- **Date de prise de fonction souhait e** : 2022-10-01
- **Dur e de contrat** : 3 ans
- **Date limite pour postuler** : 2022-09-30

Contacts

-  quipe Inria : COAST
- **Directeur de th se** :
Ignat Claudia-lavinia / claudia.ignat@inria.fr

A propos d'Inria

Inria est l'institut national de recherche d di  aux sciences et technologies du num rique. Il emploie 2600 personnes. Ses 215  quipes-projets agiles, en g n ral communes avec des partenaires acad miques, impliquent plus de 3900 scientifiques pour relever les d fis du num rique, souvent   l'interface d'autres disciplines. L'institut fait appel   de nombreux talents dans plus d'une quarantaine de m tiers diff rents. 900 personnels d'appui   la recherche et   l'innovation contribuent   faire  merger 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

- Engineering and/or Master 2 degree in Computer science / Applied mathematics with an experience in computer networks.
- Theoretical expertise: distributed systems, P2P networks
- Good collaborative and networking skills, excellent written and oral communication in English
- Good programming skills
- Strong analytical skills

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

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.