



Offer #2022-05223

PhD Position F/M Reliable and cost-efficient data placement and repair in P2P storage over immutable data

Contract type : Fixed-term contract

Level of qualifications required : Graduate degree or equivalent

Fonction : PhD Position

About the research centre or Inria department

The Inria Rennes - Bretagne Atlantique Centre is one of Inria's eight centres 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 PME's, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institute, etc.

Context

This PhD thesis will be in the context of a collaboration between HIVE and Myriads and Coast Inria teams. The Ph.D student will be located at Inria Center of the University of Rennes and will be visiting team Coast at Inria Nancy-Grand Est 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 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 PME's, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institutes, etc.

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 PME's, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institutes, etc.

Assignment

Recently, there is a growing trend toward highly distributed storage solutions by storing and sharing data across geo-distributed connected devices from the edge of the network to large scale data centres. An appealing solution – which we are exploring within the Inria-Hive collaborative framework – is utilizing the available storage and compute resources of connected devices (mobile/desktops) across the world to form a P2P storage system that provides data storage and sharing in a cost-efficient way. However, this requires to deal with several issues including node failures, node availabilities (churns), how to guarantee data availability and avoid data loss, etc.

Erasure coding (EC) has been progressively used in storage systems to provide high data availability with relatively less storage and energy cost compared to replication. For example, EC has been deployed in data analytic systems [1, 2, 3, 4] and in-memory storage systems on cached (hot) data [5]. EC can be an

ideal candidate for large scale peer-to-peer storage systems (exploits parallel read and write of data, involves large number of nodes in storing and repairing data). However, unlike previous efforts where EC is mainly used for achieved data in P2P system [6, 7], performing EC on the critical path of data access (which is the case in this project) in large scale P2P storage system (to store hot and frequently accessed data) poses many research challenges on how to ensure high data availability and meet data and node dynamicity, and on how to provide cost-effective and heterogeneity-aware data repair.

This PhD thesis will address the problem of how to provide cost-efficient yet reliable data management when deploying erasure codes (EC) in large scale trusted peer-to-peer cloud storage systems.

References:

- [1] H. Chen, H. Zhang, M. Dong, Z. Wang, Y. Xia, H. Guan, and B. Zang. "Efficient and available in- memory KV-store with hybrid erasure coding and replication". In: ACM Transactions on Storage (TOS) 13.3 (2017), pp. 1–30. doi: 10.1145/3129900.
- [2] J. Darrous and S. Ibrahim. "Understanding the performance of erasure codes in hadoop distributed file system". In: CHEOPS@EuroSys 2022: Proceedings of the Workshop on Challenges and Opportunities of Efficient and Performant Storage Systems, Rennes, France, 5 April 2022. Ed. by M. Kuhn, K. Duwe, J. Acquaviva, K. Chasapis, and J. Boukhobza. ACM, 2022, pp. 24–32. doi: 10.1145/3503646.3524296.
- [3] J. Darrous, S. Ibrahim, and C. Peñez. "Is it Time to Revisit Erasure Coding in Data-Intensive Clusters?" In: 27th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, MASCOTS 2019, Rennes, France, October 21–25, 2019. IEEE Computer Society, 2019, pp. 165–178. doi: 10.1109/MASCOTS.2019.00026.
- [4] Z. Zhang, A. Deshpande, X. Ma, E. Thereska, and D. Narayanan. Does erasure coding have a role to play in my data center? Tech. rep. MSR-TR-2010-52. May 2010.
- [5] K. V. Rashmi, M. Chowdhury, J. Kosaian, I. Stoica, and K. Ramchandran. "EC-Cache: Load-Balanced, Low-Latency Cluster Caching with Online Erasure Coding". In: 12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016, Savannah, GA, USA, November 2–4, 2016. Ed. by K. Keeton and T. Roscoe. USENIX Association, 2016, pp. 401–417.
- [6] J. Kubiawicz, D. Bindel, Y. Chen, S. Czerwinski, P. Eaton, D. Geels, R. Gummadi, S. Rhea, H. Weatherspoon, W. Weimer, C. Wells, and B. Zhao. "OceanStore: An Architecture for Global- Scale Persistent Storage". In: SIGPLAN Not. 35.11 (Nov. 2000), pp. 190–201. issn: 0362-1340. doi: 10.1145/356989.357007.
- [7] R. Rodrigues and B. Liskov. "High Availability in DHTs: Erasure Coding vs. Replication". In: Peer-to-Peer Systems IV, 4th International Workshop, IPTPS 2005, Ithaca, NY, USA, February 24–25, 2005, Revised Selected Papers. Ed. by M. Castro and R. van Renesse. Vol. 3640. Lecture Notes in Computer Science. Springer, 2005, pp. 226–239. doi: 10.1007/11558989_21.

Main activities

- As a first step, we will investigate new data placement strategies that can ensure high data availability under frequent failures and node unavailabilities. We will start by exploring how to initially place the data while considering the performance of data retrieval (location-aware, upload bandwidth of storage nodes) and the availability of data (fast data repair to avoid data loss in case of multiple failures and node unavailabilities "churns").
- Data comes with different format and has different access patterns (write once read many, periodically read, hot and cold, etc). In addition, P2P systems are highly dynamic (nodes availabilities, churns, the contributed storage and bandwidth of the nodes vary with time, etc). Therefore, data should be re-encoded (e.g., using wide-strip for cold data) and re-placed considering the dynamicity of both data and participating nodes. To facilitate that, we will make use of machine learning and probabilistic models to predicate node availabilities and study how to use the role of nodes, the contributed resources of the nodes and the incentive and awarding mechanisms which are used in the system to classify and estimate the performance of storage nodes.
- To cope with the high number of repair jobs due to lost or temporarily unavailable data, we will design data repairs orchestrator (centralized/decentralized scheduling framework) that ensures cost-effective and efficient data repair jobs by considering the location of data, availabilities of nodes, heterogeneity of the network bandwidth, etc.

Skills

- 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

Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Possibility of teleworking (90 days per year) and flexible organization of working hours
- partial payment of insurance costs

Remuneration

Monthly gross salary amounting to :

- 1982 euros for the first and second years and
- 2085 euros for the third year

General Information

- **Theme/Domain** : Distributed Systems and middleware System & Networks (BAP E)
- **Town/city** : Rennes
- **Inria Center** : [Centre Inria de l'Université de Rennes](#)
- **Starting date** : 2023-10-01
- **Duration of contract** : 3 years
- **Deadline to apply** : 2023-09-16

Contacts

- **Inria Team** : [MYRIADS](#)
- **PhD Supervisor** :
Ibrahim Shadi / Shadi.Ibrahim@inria.fr

About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

Warning : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

Instruction to apply

Please submit online : your resume, cover letter and letters of recommendation eventually

For more information, please contact shadi.ibrahim@inria.fr

Defence Security :

This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy :

As part of its diversity policy, all Inria positions are accessible to people with disabilities.