

**Mercredi 4 Mai 2022****14h00****UT3 Paul Sabatier, IRIT, Salle des Thèses****Morgan SEGUELA****Equipe PYRAMIDE, Equipe SEPIA, IRIT****Stratégie de réplication de données dans des systèmes larges échelles pour prendre en compte la consommation énergétique et la dépense***Jury :*

- M. Jean-Marc PIERSON – Université Toulouse III - Paul Sabatier, Directeur de thèse
- M. Riad MOKADEM – Université Toulouse III - Paul Sabatier, Co-directeur de thèse
- M. Sébastien MONNET – Université Savoie Mont Blanc, Rapporteur
- M. Nouredine MELAB – Université de Lille, Rapporteur
- Mme Esther PACITTI – Université de Montpellier, Examinatrice
- Mme Francine KRIEF – INP Bordeaux - LABRI, Examinatrice
- Mme Patricia STOLF – Université Toulouse Jean Jaurès - UT2J, Examinatrice

**Mots-clés :** Systèmes Cloud, Réplication de données, Modèle de coût économique, Consommation énergétique, Service Level Agreement

**Résumé :** Les applications d'aujourd'hui nécessitent l'accès à des données réparties à travers le monde. De plus, l'augmentation du volume de ces données conduit à des problématiques de disponibilité et de performance, surtout lorsque ces données sont fréquemment requêtées. Une manière de répondre à ces problèmes est la réplication de données, une technique très utilisée dans les systèmes distribués classiques, mais également dans les systèmes à grande échelle. De nombreuses stratégies de réplication de données ont été proposées dans de tels systèmes. Elles visent à déterminer les données à répliquer, combien de répliques créer, quand créer les répliques et où les placer. Dans ce manuscrit, nous proposons une stratégie de réplication de données dans les systèmes Cloud. De nos jours, les enjeux environnementaux deviennent des problématiques majeures dans notre société. Certaines entreprises cherchent à avoir un impact positif sur ces enjeux. Pour les fournisseurs de services numériques et de Cloud, cela se traduit par une réduction de l'émission de gaz à effet de serre en réduisant la consommation énergétique et en utilisant des ressources énergétiques plus vertes. Dans le cadre de cette thèse, nous proposons une stratégie de réplication de données qui prend en compte ces problématiques économiques et énergétiques. Un placement initial, statique, est mis en place pour répondre aux objectifs de disponibilité et de tolérance aux fautes tout en tenant compte de la consommation énergétique et de la dépense. Ce placement s'appuie sur l'hétérogénéité entre les centres de données et sur l'utilisation de technique de veille permettant de réduire la consommation énergétique de serveurs inutilisés. Un tel placement permet de choisir la politique à mettre en place selon l'intérêt du fournisseur. Une gestion dynamique des répliques, s'appuyant sur le placement initial, est ensuite proposée. Elle permet de s'adapter aux variations de la charge de travail, détectée à l'aide de Cartes de contrôle qui s'appuient sur des probabilités pour lever des alertes. Ainsi, de nouvelles répliques sont créées lorsque la charge augmente, tout en tenant compte de la consommation énergétique. D'un autre côté, des répliques sont supprimées lorsque cette



charge diminue. Cela permet de réduire les coûts et la consommation énergétique liés au stockage. Une évaluation de performances, via une simulation, a permis de valider la stratégie proposée tout en comparant ses performances à celles d'autres stratégies proposées dans la littérature. Pour le placement initial, ces évaluations ont permis de mettre en avant l'impact des différents choix proposés, mais aussi de montrer qu'un placement initial des répliques plus "intelligent" peut avoir un impact positif sur les performances et la disponibilité, tout en réduisant la consommation énergétique et les dépenses. Puis, nous l'avons évalué en intégrant la gestion dynamique des répliques. Les résultats de la partie dynamique montrent que nous arrivons à améliorer les performances par rapport au placement initial, tout en réduisant à la fois les dépenses et la consommation énergétique. De plus, cela permet de mettre en avant l'importance des choix et des méthodes mises en place lors de l'ajout ou de la suppression des répliques.

