Université de Bretagne Occidentale

Ecole Doctorale

Mathématiques et Sciences et Technologies de l'Information et de la Communication

ENSTA Bretagne/Lab-STICC (UMR 6285)

AVIS DE SOUTENANCE DE THESE EN COTUTELLE

avec l'USTHB - ALGER (Algérie)

Le lundi 18 avril 2022 à 10h (heure locale)

à l'USTHB, Faculté d'Informatique, salle de réunion, ALGER.

Madame CHIKHAOUI AMINA

soutiendra une thèse de doctorat sur le sujet suivant :

" Vers une approche intelligente de placement de données dans un cloud distribué basé sur un système de stockage hybride ".

Le jury sera ainsi composé :

- M. BARAIS OLIVIER, Professeur des universités
IRISA - INRIA - RENNES

- M. BOUKHALFA KAMEL, Professeur
USTHB - ALGER - ALGERIE

- M. BOUKHOBZA JALIL, Professeur des universités
ENSTA Bretagne - BREST

- MME NASSER HASSINA, Professeure
USTHB - ALGER - ALGERIE

- M. ZITOUNI ABDELHAFID, Professeur
Université A. Mehri Constantine - CONSTANTINE - ALGERIE

invité(e) :

- M. LEMARCHAND LAURENT, Maître de conférences
Univ. de Bretagne Occidentale - BREST

A BREST, le 05 avril 2022

Le Président de l'Université de Bretagne Occidentale,

UNIVERSITÉ DE BRETAGNE OCCIDENTALE

M. GALLOU

Résumé de thèse de Doctorat de CHIKHAOUI Amina

Titre: Vers une approche intelligente de placement de données dans un *Cloud* distribué basé sur un système de stockage hybride
Directeurs de thèse: BOUKHALFA Kamel et BOUKHOBZA Jalil

La fédération de *Clouds* permet d'étendre de manière transparente les ressources des fournisseurs de services *Cloud* (Cloud service Providers: CSP) afin de fournir une meilleure qualité de service (Quality of Service: QoS) aux clients sans frais de déploiement supplémentaires.

Le stockage en tant que service (Storage as a Service: StaaS), constitue l'un des principaux services *Cloud* offerts aux clients. Pour un tel service, la performance des Entrées/Sorties (E/S) des supports de stockage et la latence du réseau sont parmi les métriques les plus importantes considérées par les clients. En effet, le système d'E/S prend environ 90% du temps d'exécution d'une transaction pour certaines requêtes de base de données. Afin de rassurer les clients, certaines sociétés de *Cloud* incluent déjà des garanties de latence dans leurs contrats de niveau de service (Service Level Agreements: SLA) et les clients peuvent payer des frais supplémentaires pour réduire davantage les latences.

Cette thèse traite du problème de placement de données pour un CSP faisant partie d'une fédération. En effet, offrir des services attractifs et peu coûteux est un grand défi pour les CSP. Notre but est de proposer des approches intelligentes pour un meilleur placement des données qui minimise le coût pour le fournisseur tout en satisfaisant les clients. Cette approche doit prendre en compte l'hétérogénéité des ressources de stockage interne et externe en terme de plusieurs paramètres (comme la capacité, les performances, la tarification) ainsi que les caractéristiques des clients et leurs exigences. En dépit du fait que de nombreuses stratégies de placement de données ont été proposées pour les systèmes de stockage hybrides, elles ne sont pas généralisables pour toutes les architectures. En effet, une stratégie de placement doit être conçue selon l'architecture du système pour laquelle elle est proposée et les objectifs visés.

Les stratégies d'optimisation dans les environnements *Cloud* sont, de manière générale, basées sur le coût. En effet, réduire les coûts opérationnels tout en maintenant des niveaux élevés de garantie de service pour les clients sont des facteurs importants pour que les fournisseurs augmentent leurs revenus et restent compétitifs. Ainsi, notre première contribution est relative à l'évaluation du coût de placement de données dans un *Cloud* fédéré. Le modèle proposé prend en compte différents facteurs modélisant les caractéristiques physiques et fonctionnelles du système fédéré.

Afin d'optimiser le placement de données, les coûts de stockage, de migration et de latence doivent être pris en compte. Ces coûts sont corrélés et s'avèrent, dans certains cas, contradictoires. Il convient alors de trouver un compromis entre ces coûts. Par conséquent, une optimisation multi-objectifs est nécessaire. Dans cette perspective, notre deuxième contribution consiste en la proposition d'une approche multi-objectifs basée sur la méta-heuristique évolutionnaire NSGAII. L'approche développée CDP-NSGAII$_{IR}$ *(a Constraint Data Placement matheuristic based on NSGAII with Injection and Repair functions)* est une matheuristique de placement de données qui consiste à hybrider une méthode exacte avec la méta-heuristique NSGAII. Un opérateur de réparation a été conçu pour rendre les solutions réalisables au regard des contraintes du système.

Enfin, la troisième contribution de ce travail de thèse concerne la réplication de données. En effet, fournir des services fiables avec une haute disponibilité des données et des performances

adéquates sont des exigences clés qui doivent être satisfaites. Le concept de réplication est utilisé pour garantir de telles exigences. Pour résoudre le problème de placement et de sélection des répliques, nous avons proposé *StorNIR* (*a cost-efficient data object **Stor**ing scheme based on NSGAII upgraded with **I**njection and **R**eparation operators*), une stratégie de stockage d'objets basée sur NSGAII et mise à niveau avec des opérateurs d'injection et de réparation. Le même raisonnement de la deuxième contribution a été repris dans ce travail. En effet, nous avons utilisé la matheuristique combinant une méthode exacte avec NSGAII pour résoudre le problème de placement et de sélection des répliques. Il est à noter que nous avons mis à jour les opérateurs d'évolution de NSGAII pour les adapter au problème de réplication. De plus, lorsque les objets sont répliqués, la latence réseau des requêtes d'accès peut être optimisée. Par conséquent, dans l'approche matheuristique nous avons proposé un algorithme pour optimiser la latence des requêtes. Il consiste à affecter les requêtes à l'emplacement le plus proche qui satisfait les performances E/S de stockage.

Mots clés— Informatique en nuages,Fédération de *Clouds*, Placement de données, Réplication de données, Optimization, NSGAII, Matheuristique.