

Elasticité à approche combinée pour le traitement des flux dans les environnements de brouillard

Alessio Pagliari, Guillaume Pierre

Univ Rennes, Inria, CNRS, IRISA
alessio.pagliari@irisa.fr

Résumé

Le traitement des données en temps réel est une exigence standard du Fog Computing. L'adaptation dynamique [3] des plateformes de data streaming est une fonctionnalité essentielle pour leur permettre de gérer efficacement les charges de travail variables dans le temps et d'optimiser la consommation des ressources. Cependant, la mise à l'échelle horizontale seule se confronte à des limites lorsque les ressources de calcul disponibles sont limitées, que les temps de reconfiguration de la plateforme de streaming sont importants ou que, en raison des exigences contextuelles, l'application ne peut pas être mise à l'échelle horizontalement. Dans de telles situations, le calcul approximatif [5] peut être une alternative en permettant aux applications d'échanger la précision du résultat qu'elles produisent pour accélérer le traitement. Cependant, les applications critiques peuvent nécessiter une haute précision, ou des tâches spécifiques ne peuvent pas supporter l'approximation par conception.

Nous trouvons nécessaire un système d'adaptation qui combine plusieurs techniques, telles que l'adaptation des ressources et le calcul approximatif. Pour contrôler l'approximation avec précision, nous exploitons le paradigme de l'informatique de transprécision [4]. Une nouvelle approche qui envisage le contrôle de l'approximation sur le "quand" et le "où" est appliquée en changeant dynamiquement le calcul d'une tâche pour s'adapter aux changements du temps d'exécution.

Nous proposons TransScale, un auto-scaler à approche combinée qui combine la mise à l'échelle horizontale [2] et le contrôle de la transprécision pour surmonter les limites des environnements fog. Nous concevons TransScale de manière à rendre la méthode d'approximation transparente pour le système et à la prendre en charge quelle que soit l'implémentation définie par l'utilisateur. Nous mettons en œuvre un composant de prédiction des performances [1] combiné pour optimiser la reconfiguration en tenant compte de l'impact des deux approches sur les performances. Nous définissons cinq stratégies de mise à l'échelle basées sur les exigences de qualité de service pour décider de la meilleure reconfiguration, chacune visant à optimiser une métrique différente.

En fonction de l'objectif de la politique de reconfiguration, nous montrons que TransScale peut réduire les occurrences de reconfiguration, optimiser l'utilisation des ressources et soutenir des charges de travail élevées dans des environnements à ressources limitées.

Mots-clés : Traitement des flux de données, élasticité, calcul transprécis, informatique en brouillard.

Bibliographie

1. Arkian (H.) et al. – An experiment-driven performance model of stream processing operators in fog computing environments. – In *Proc. ACM SAC*, 2020.
2. Arkian (H.) et al. – Model-based stream processing auto-scaling in geo-distributed environments. – In *Proc. ICCCN*, 2021.
3. Cardellini (V.) et al. – Run-time adaptation of data stream processing systems : The state of the art. *ACM CSUR*, 2022.
4. Malossi (A. C. I.) et al. – The transprecision computing paradigm : Concept, design, and applications. – In *Proc. DATE*, 2018.
5. Quoc (D. L.) et al. – StreamApprox : Approximate computing for stream analytics. – In *Proc. ACM/IFIP/USENIX Middleware*, 2017.