

Simulation d'algorithmes d'équilibrage de charge dans un environnement distribué

Éléments bibliographiques

Kévin Barreau Mounir Hamdane Guillaume Marques
Corentin Salingue

20 janvier 2015

Rappel du contexte

Nous devons développer une solution logicielle permettant de tester des algorithmes d'équilibrage de charges dont le principe est indiqué par nos clients. Nous devons donc les implémenter.

L'environnement de simulation voulu est un système distribué constitué de n noeuds de stockage dans lequel on souhaite stocker m objets. Un générateur de requêtes, que nous devrons implémenter, enverra des messages au système pour consulter un certain nombre d'objets.

Le but est de déterminer quel algorithme distribue le plus efficacement les requêtes, afin d'obtenir le délai le moins long entre l'arrivée d'une requête dans le système et son traitement.

Le client souhaite aussi avoir à sa disposition un outil pour visualiser l'état du système en temps réel. Par exemple, la charge (nombre de requêtes en attente) de chaque noeud.

1 Système distribué

- Le livre [ÖV11] apporte les principes sur les bases de données distribuées.

2 Algorithmes d'équilibrage de charges

- L'article [ABKU99] met en exergue des algorithmes d'équilibrage, sources d'inspiration pour les clients dans l'élaboration de leurs algorithmes.
- Le papier [MRS05], de la même manière que le précédent, s'intéresse de plus près à des algorithmes d'allocation aléatoire.
- L'article [FKSS14] est aussi important car il traite des algorithmes d'équilibrage des charges dans un réseau pair à pair.

3 Cassandra

- Le livre [Hew10] renseigne sur la compréhension du fonctionnement de Cassandra, ainsi que les structures particulières de ce système de base de données.
- Le lien [Fou09] permet d'accéder à l'ensemble de la documentation technique et au dépôt des sources de la base de données Cassandra.
- Le papier [LA09] apporte les explications des créateurs de Cassandra.
- Le lien [Dat15] possède la documentation la plus complète sur la dernière version de Cassandra, et sur les manières de l'utiliser. Les nombreux exemples et les approfondissements sont des points d'appui importants pour le projet.

4 Visualisation

- Le lien [Dat14] est intéressant pour le projet car il apporte un logiciel capable de montrer graphiquement des statistiques sur une base de données Cassandra opérationnelle. Il peut être utilisé dans le projet, ou apporté des pistes pour la visualisation des données.
- L'article [AAB⁺12] s'intéresse à un logiciel de représentation de graphe développé au Labri. La proximité des créateurs peut apporter une solution viable.

Références

- [AAB⁺12] D. Auber, D. Archambault, R. Bourqui, A. Lambert, M. Mathiaut, P. Mary, M. Delest, J. Dubois, and G. Melançon. The tulip 3 framework : A scalable software library for information visualization applications based on relational data. [Research Report] RR-7860, hal-00659880, 2012.
- [ABKU99] Y. Azar, A. Broder, A. Karlin, and E. Upfal. Balanced allocations. *SIAM Journal on Computing*, 29(1) :180–200, 1999.
- [Dat14] DataStax. Datastax opscenter : Datastax. <<http://www.datastax.com/what-we-offer/products-services/datastax-opscenter>>, 2014. [Accessed 15 January 2015].
- [Dat15] DataStax. Datastax cassandra 2.1 documentation. <<http://www.datastax.com/documentation/cassandra/2.1/cassandra/gettingStartedCassandraIntro.html>>, 2015. [Accessed 15 January 2015].
- [FKSS14] P. Felber, P. Kropf, E. Schiller, and S. Serbu. Survey on load balancing in peer-to-peer distributed hash tables. *IEEE Communications Surveys and Tutorials*, 16(1) :473–492, 2014.
- [Fou09] The Apache Software Foundation. The apache cassandra project. <<http://cassandra.apache.org>>, 2009. [Accessed 15 January 2015].
- [Hew10] E. Hewitt. *Cassandra : The Definitive Guide*. O’Reilly Media, 2010.
- [LA09] M. Prashant L. Avinash. Cassandra - a decentralized structured storage system. 2009.
- [MRS05] M. Mitzenmacher, A. Richa, and R. Sitaraman. Tthe power of two random choices : A survey of techniques and results. 2005.
- [ÖV11] T. Özsu and P. Valduriez. *Principles of Distributed Database Systems*. Computer science. Springer, 2011.