

Sujet de master recherche « Architectures logicielles distribuées » 2006–2007

## Regroupements sémantiques de pairs

Encadrant principal : Sylvie CAZALENS  
courriel : [Sylvie.Cazalens@lina.univ-nantes.fr](mailto:Sylvie.Cazalens@lina.univ-nantes.fr)  
tél. : 02 51 12 58 45

Co-encadrant(s) : Philippe LAMARRE  
courriel : [Philippe.Lamarre@lina.univ-nantes.fr](mailto:Philippe.Lamarre@lina.univ-nantes.fr)  
tél. : 02 51 12 58 50

**Prérequis** : Java

**Mots-clés** : Système pair à pair (P2P), Organisation sémantique, Routage de requête.

## Objectif du stage

Les systèmes Pair à Pair (P2P) offrent un moyen, de plus en plus reconnu, de partager des informations facilement, tout en assurant l'autonomie et la vie privée des pairs.

APPA (*Atlas Peer-to-Peer Architecture*) [1] est un nouveau système de gestion de données P2P que nous construisons au LINA. Les objectifs d'APPA sont le passage à l'échelle, la disponibilité des données et les performances pour des applications avancées. APPA a une architecture indépendante du réseau, en termes de services de base et de service avancés, afin de pouvoir fonctionner sur différents réseaux P2P (non structuré, DHT, super-peer). Les services avancés sont le traitement de requêtes, la réplication de données, et la gestion de cache. APPA est en cours de réalisation avec le framework JXTA<sup>1</sup>.

Lorsque le nombre de pairs est grand, le routage des requêtes dans de tels systèmes peut s'avérer inefficace. C'est le cas lorsque les requêtes passent d'un pair à un autre choisi aléatoirement. Une façon d'améliorer le routage d'une requête est de regrouper les pairs. Cela permet de mieux cibler les pairs auxquels elle est envoyée. Dans une telle organisation, la requête peut circuler de groupe en groupe, jusqu'à atteindre le ou les groupes les plus pertinents où elle peut être traitée.

Le stage concerne le regroupement sémantique de pairs : les pairs d'un même groupe partagent un certain nombre de concepts et de relations entre ces concepts (notion d'ontologie). Les informations dont ils disposent sont relatives à ces concepts ou à des concepts liés.

L'objectif du stage concerne l'étude et la simulation de différentes organisations du type décrit ci-dessus. Plus particulièrement, une étude comparative des performances pour le routage des requêtes doit être menée. Les protocoles devront être conçus pour la plateforme APPA. Des simulations seront faites avec la plate-forme de simulation PeerSim<sup>2</sup>.

---

<sup>1</sup><http://www.jxta.org>

<sup>2</sup><http://peersim.sourceforge.net/>

## Travail à réaliser

- Définition de types d'organisation
- Définition, implémentation, évaluation d'algorithmes et protocoles pour le routage des requêtes dans ces organisations.

## Références

- [1] Reza Akbarinia, Vidal Martin, Esthere Pacitty, and Patrick Valduriez. Replication and Query Processing in the APPA Data Management System. In *Distributed Data& structures (WDAS), records of the 6th International Meeting (Lausanne, Switzerland), Waterloo*, 2004.
- [2] Anne Doucet and Nicolas Lumineau. A collaborative approach for query propagation in peer to peer systems. In *International Workshop on Semantic Web and Database*, 2003.
- [3] Manolis Koubarakis, Peter Triantafillou, Chyssani Xiruhaki, and Nikolas Ntarmos. Toward high performance peer to peer content and ressource sharing systems. In *Conference on Innovative Data Systems Research*, 2003.
- [4] Arturo Crespo and Hector Garci-Molina. Semantic overlay network for peer to peer systems. Technical report, Computer Science Department of Stanford University, october 2002.