

[illegible]

Ibrahima NLANG

Dépt. Mathématiques et informatique
Université Cheikh Anta Diop de Dakar
BP 5005 - Dakar-Fann
SENEGAL
iniang@ucad.sn

[illegible]

ABSTRACT. This paper focuses on fault tolerance of super-nodes in P2P-SIP systems. These systems are characterized by high volatility of super-nodes. Most fault-tolerant proposed solutions are only for physical defects. They do not take into consideration the timing faults that are very important for multimedia applications such as telephony. This paper proposes a timing and physical fault tolerant mechanism based on P2P overlay with two levels for P2P-SIP systems. The simulation results show that our proposition reduces mostly the nodes location latency and increases the probability to find the called nodes.

MOTS-CLÉS : VoIP, P2P, SIP, P2P-SIP, DHT, super- nœuds, tolérance aux pannes.

KEYWORDS: VoIP, P2P, SIP, P2P-SIP, DHT, super-nodes, fault tolerance.

[illegible][illegible]

1. Introduction

Les réseaux P2P sont de trois types: P2P avec serveur central, P2P non structuré et P2P structuré. Le réseau P2P non structuré est basé sur un graphe aléatoire et utilise des inondations [10], mais ne passe pas à l'échelle à cause de ces inondations, générant des charges lourdes dans le réseau. Le réseau P2P structuré est basé sur l'utilisation de la DHT (Distributed Hash Table) qui est une grande table de hachage maintenue par les nœuds du réseau distribué. Cette méthode est plus efficace que les inondations et a une meilleure évolutivité [5]. Récemment, des travaux ont été proposés pour intégrer le système P2P et les services de téléphonie sur IP avec l'utilisation du protocole SIP (P2P-SIP). Avec la forte volatilité des nœuds, des solutions de gestion de la tolérance aux pannes ont été proposées mais ces approches sont très limitées pour prendre en compte ce nouveau service de voix sur IP.

Dans cet article, nous proposons une solution efficace de tolérance aux pannes des super-nœuds dans les systèmes de VoIP P2P-SIP. Notre approche s'appuie sur un schéma DHT hiérarchique à trois niveaux. L'architecture proposée améliore la gestion des pannes physiques existante et introduit une solution pour gérer les pannes temporelles. Le reste de cet article est organisé comme suit. La section 2 présente les problématiques de recherche. La section 3 présente notre schéma DHT hiérarchique. Une validation expérimentale et théorique est proposée dans la section 4. Enfin, une conclusion est formulée dans la section 5.

2. Problématiques de recherche

2.1. Structuration hiérarchique dans les systèmes P2P

La plupart des schémas DHT existants sont basés sur une architecture plate. Toutefois, des solutions de DHTs hiérarchiques ont été proposées visant à améliorer l'évolutivité et la latence. Dans [6], Coral, un anneau Chord hiérarchisé à trois niveaux, est présenté. Cette solution est un système pair à pair de distribution de contenu. Coral permet aux nœuds de localiser les nœuds et de télécharger des fichiers à partir de leur nom. Dans [7], il a été proposé un réseau hiérarchique, nommé HP2P, qui est un réseau P2P hiérarchique et hybride par rapport à d'autres. Le réseau HP2P associe la DHT et les méthodes d'inondation dans une topologie complètement distribuée. Dans HP2P, Chord est utilisé pour la couche supérieure, alors que l'inondation est utilisée pour la couche inférieure. En outre, dans la couche supérieure, chaque nœud est appelé un nœud virtuel car il représente un groupe de nœuds de la couche inférieure.

2.2. Tolérance aux pannes dans les réseaux P2P

La quasi-totalité des travaux concernant la tolérance aux pannes dans les systèmes P2P traitent les partages de fichiers mais pas des services de voix sur IP [2, 3, 4]. Dans [10], la méthode de tolérance aux pannes des super pairs proposée est basée sur la technique de redondance. Dans chaque groupe, les k super-pairs sont spécifiés pour former un super-pair virtuel. Puis, le super-pair virtuel sert les pairs réguliers au sein de son groupe. L'article [3] présente une approche similaire à [4]. Dans cette architecture hiérarchique, si le super-pair échoue, ces pairs réguliers ne pourront plus répondre aux requêtes.

2.3. Problématiques de recherche sur la tolérance aux pannes dans P2P-SIP

Tout système P2P tolérant aux pannes de nœud comprend un mécanisme chargé de détecter les défaillances, et de les gérer d'une manière transparente. Dans les réseaux P2P-SIP [1, 9], jusqu'alors proposés, ce mécanisme de détection ne fait pas de différence entre les défaillances physiques et les défaillances temporelles. Si, après un certain délai de garde, un super-nœud n'envoie pas de messages de rafraîchissement à ses successeurs ou d'accusés de réception à ses nœuds ordinaires, il sera alors considéré comme étant en panne. En plus, les solutions proposées ne prennent pas entièrement en charge toutes les défaillances des super-nœuds. Il n'existe pas de mécanisme de reprise automatique après la panne de super-nœuds d'attache car ce type de nœud est unique pour un nœud ordinaire bien déterminé. En cas de panne de son super-nœud d'attache, un nœud ordinaire ne peut ni passer ni recevoir un appel jusqu'au prochain rafraîchissement. Une autre limite de ces solutions est la non prise en compte de la défaillance temporelle des super-nœuds qui constitue un aspect très important pour une application multimédia comme la téléphonie.

3. Schéma DHT hiérarchique pour la tolérance aux pannes dans les réseaux P2P-SIP

3.1. Architecture de la solution proposée

Notre solution est basée sur une architecture hiérarchique à trois niveaux contrairement à l'architecture de P2P-SIP classique qui repose sur deux niveaux (voir Figure 1). Les deux premiers niveaux sont les mêmes que ceux décrits dans l'architecture P2P-SIP classique, c'est-à-dire un premier niveau pour les nœuds

ordinaires et un deuxième niveau pour les super-nœuds. Nous proposons un troisième niveau pour les super-nœuds complexes. Un super-nœud complexe est un nœud stockant les enregistrements de certains nœuds ordinaires. Il participe au routage des messages de localisation en cas de certaines défaillances de super-nœuds. Les super-nœuds complexes sont sélectionnés parmi les super-nœuds. Le choix est basé sur les paramètres suivants : la durée de vie du nœud dans le réseau, la bande passante, la vitesse du CPU, la taille de la mémoire. Les super-nœuds complexes sont proposés d'une part pour apporter une réponse à la question de la tolérance aux défaillances physiques des super-nœuds d'attache, et d'autre part pour introduire la gestion de la tolérance liée aux défaillances temporelles de tous les types de super-nœuds. Dans notre solution, chaque nœud ordinaire est à la fois rattaché à un super-nœud et à un super-nœud complexe pour créer une certaine redondance. Le choix de deux super-nœuds, dans notre proposition, est guidé par le fait que chaque nœud ordinaire doit rafraîchir son enregistrement auprès de ses nœuds d'attache, après un certain temps. Par conséquent, le nombre de messages de rafraîchissement sera multiplié par le nombre de nœuds d'attache pour chaque nœud ordinaire.

Nous utilisons Chord dans notre proposition parce qu'il comporte un service pair à pair évolutif de recherche de ressources pour les applications Internet.

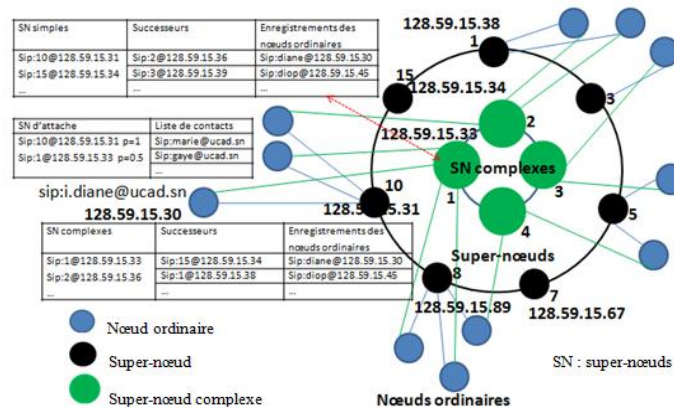


Figure 1. Architecture de la tolérance aux pannes dans les réseaux P2P-SIP

3.2. Gestion de la tolérance aux pannes des super-nœuds

Notre système fonctionne comme suit: au démarrage, un nœud ordinaire tente de se connecter à un super-nœud et à un super-nœud complexe. Pour cela, le nœud ordinaire cherche à découvrir un super-nœud qui lui servira de bootstrap (passerelle). Après cette découverte, le nœud ordinaire va essayer de s'enregistrer auprès de son super-nœud responsable via le bootstrap. Avant de confirmer l'enregistrement, le super-nœud

responsable va répliquer l'enregistrement du nœud ordinaire sur un super-nœud complexe. Après la phase d'enregistrement, le super-nœud responsable va répliquer l'enregistrement du nœud ordinaire sur ces successeurs.

3.2.1. Gestion des pannes physiques

a) Détection: Les défaillances physiques sont détectées en utilisant les messages de rafraîchissement. Ainsi, un nœud ordinaire détecte la panne de son super-nœud d'attache s'il ne reçoit pas de réponse à son rafraîchissement. Dans ce cas, le nœud ordinaire va, à nouveau, reprendre la procédure de découverte de super-nœud. Ceci n'est pas adapté pour des applications qui ont des contraintes de temps telles que la VoIP où le délai de localisation ne doit pas être important.

b) Recouvrement: Trois cas de figure sont considérés : i) Si le super-nœud défaillant est le nœud responsable de la clé du nœud ordinaire local, alors le système fait recourt à un super-nœud complexe. Le nœud ordinaire envoie à nouveau son message à son super-nœud complexe. Ce dernier envoie le message au super-nœud complexe du nœud ordinaire distant qui va traiter le message et renvoyer les informations de localisation du nœud ordinaire recherché. ii) Si le super-nœud défaillant est un successeur du nœud responsable de la clé du nœud ordinaire local, alors il sera ignoré au profit de son successeur immédiat. iii) Si le super-nœud défaillant est le nœud responsable de la clé du nœud ordinaire distant, alors son successeur immédiat se chargera de traiter le message et d'envoyer les informations de localisation du nœud ordinaire.

3.2.2. Gestion des pannes temporelles

a) Détection : A la différence du mécanisme utilisé dans la gestion des pannes physiques, nous utilisons ici la technique de messages "ping/pong" pour détecter les défaillances temporelles. Pour cela, de manière périodique, un nœud envoie un message "ping" à un autre. Cette technique permet une détection plus ciblée car chaque message "ping" réclame explicitement un message "pong".

b) Recouvrement: Après avoir détectée une panne, la prochaine étape consiste à la recouvrir. La principale technique pour atteindre cet objectif est de répliquer les enregistrements. Nous considérons aussi trois scénarios pour la gestion des pannes temporelles: i) Si le super-nœud défaillant est le responsable de la clé du nœud ordinaire, alors le nœud ordinaire annule le message envoyé au super-nœud défaillant et renvoie le même message à nouveau à son super-nœud complexe. Ce dernier va traiter le message, c'est-à-dire, envoyer le message au super-nœud complexe du nœud ordinaire distant. ii) Si le super-nœud défaillant est un successeur du super-nœud responsable de la clé du nœud ordinaire, alors le message est envoyé à un super-nœud complexe connu par le super-nœud responsable. Ce dernier envoie le message au super-nœud complexe du nœud ordinaire distant. iii) Si le super-nœud défaillant est le responsable de la clé du

nœud ordinaire distant, alors le message est envoyé à un super-nœud complexe connu par le prédécesseur du super-nœud défaillant. Ce super-nœud complexe envoie le message au super-nœud complexe du nœud ordinaire distant

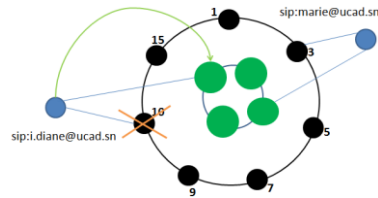


Figure 2. Gestion de la défaillance temporelle du super-nœud responsable

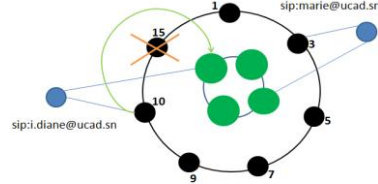


Figure 3. Gestion de la défaillance temporelle d'un successeur du super-nœud responsable

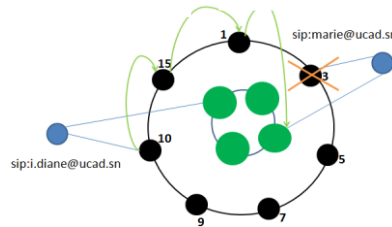


Figure 4. Gestion de la défaillance temporelle du super-nœud responsable distant

4. Validation expérimentale et théorique

Cette section présente les travaux relatifs à la validation expérimentale de notre proposition. Durant les expériences, nous avons utilisé le logiciel de simulation de réseau pair-à-pair appelé OverSim [8]. Nous fixons le pourcentage de pannes à 70%, faisons varier le nombre de super-nœuds de 100 à 2000 et mesurons à chaque fois la latence de localisation et la probabilité de localiser un nœud ordinaire.

4.1. Notre solution Vs. P2P-SIP Classique

Dans cette sous section, nous comparons notre proposition et le système P2P-SIP classique par rapport à la latence et à la probabilité de localiser un nœud ordinaire.

a) Latence de localisation: Les résultats montrent qu'il y a, quelque soit le type de panne considéré, une nette différence entre le système P2P-SIP classique et notre proposition vis-à-vis de la latence de localisation des nœuds ordinaires.

b) Ratio de localisation: Cette comparaison est faite par rapport à la probabilité de trouver un nœud ordinaire.

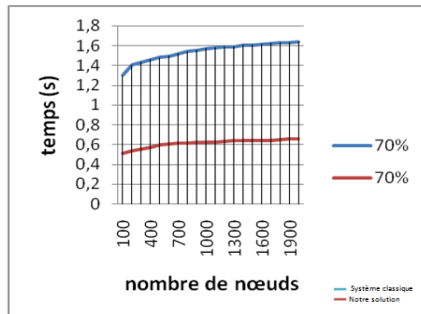


Figure 5. Latence pour 70% de pannes temporelles

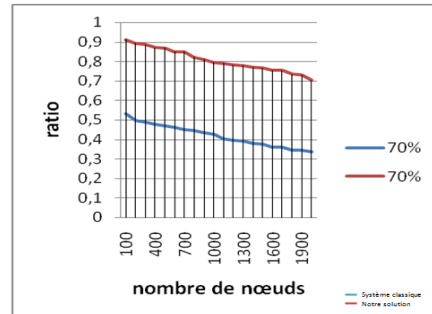


Figure 6. Ratio pour 70% de pannes temporelles

4.2. Surcharge du réseau

Cette sous section est une comparaison de notre proposition et du système P2P-SIP classique par rapport à la surcharge du réseau. Nous considérons un réseau avec n nœuds ordinaires, n' super-nœuds et m le nombre de répliques des enregistrements.

	P2P-SIP classique	Notre proposition
Nombre de messages de rafraîchissement	$n \times m$	$n \times m$
Nombre de messages "ping/pong"	0	$n + n' \times m$

Figure 7. Surcharge du réseau : P2P-SIP classique vs notre proposition

5. Conclusion

Dans ce papier, nous avons proposé une approche hiérarchique pour la gestion de la tolérance aux pannes dans les réseaux P2P-SIP. Cette approche combine la gestion des pannes physiques et temporelles. Elle est basée sur l'architecture P2P-SIP à trois niveaux qui correspondent aux trois types de nœuds (nœud ordinaire, super-nœud et super-nœud complexe). Notre proposition permet d'améliorer considérablement les performances de la téléphonie P2P-SIP. Toutefois, il est important de noter que notre approche est plus coûteuse en termes de nombre de messages. En plus des messages de rafraîchissement, nous utilisons des messages «ping / pong» pour détecter les pannes temporelles. Cette approche est meilleure vis à vis de la stabilité, de l'évolutivité et de l'efficacité du système par rapport aux autres réseaux.

6. Bibliographie et biographie

6.1 Bibliographie

- [1] Bryan, D., B. Lowekamp, and C. Jennings “*SOSIMPLE: A Serverless, Standard-based, P2P SIP Communication System*,” International Workshop on Advanced Architectures and Algorithms for Internet Delivery and Applications, June 2005.
- [2] Laban Mwansa, Jan Janeček, “*Ensuring fault-tolerance in generic network location service*”, Proceedings on 22nd European conference on Modelling and simulation, 2008.
- [3] Jenn-Wei Lin, Ming-Feng Yang, Jichiang Tsai, “*Fault Tolerance for Super-peers of P2P Systems*”, Proceedings of the 13th Pacific Rim International Symposium on Dependable Computing, 2008.
- [4] B. Yang and H. Garcia-Molina, “*Designing a Super-Peer Network*”, Proc. 19th Int’l Conf. Data Eng., Mar. 2003. pp. 49-62.
- [5] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker, “*A scalable content-addressable network*”, In Proc. ACM SIGCOMM, August 2001.
- [6] Freedman, M., Mazieres, D., “*Sloppy Hashing and Self-Organising Clusters*”, 2nd International Workshop on Peer-to-Peer Systems, Berkley, USA, February 2003.
- [7] Peng, Z., Duan, Z., Qi, J., Cao, Y., Lv, E., “*HP2P: A Hybrid Hierarchical P2P Network*”, 1st International Conference on the Digital Society (ICDS), Guadeloupe, French Caribbean, January 2007.
- [8] Ingmar Baumgart, Bernhard Heep, Stephan Krause, “*OverSim: A Flexible Overlay Network Simulation Framework*”, IEEE Global Internet Symposium, 2007
- [9] Singh, K., and H. Schulzrinne, “*Peer-to-Peer Internet Technology Using SIP*,” Columbia University Technical Report CUCS-044-04, New York, October 2004
- [10] G.X. Yue, R.F. Li, and Z.D. Zhou, “*A P2P network model with multi-layer architecture*”, based on region, Journal of Software, 2005,16(6):1140.1150. DOI: 10.1360/jos161140, 2005.

6.2 Biographie

Ibrahima DIANE : Etudiant en thèse à l'école doctorale de Maths-Info de l'Université C. A. Diop de Dakar. Domaine de recherche : QoS dans les réseaux P2P.

Ibrahima NIANG : Dr. En informatique au dépt de Maths-Info à l'Université Cheikh Anta Diop de Dakar. Domaine de recherche : QoS, Sécurité et Mobilité dans les P2P et Sans fils.