LIGNON Rodolphe et LE NORCY Arnaud
Encadrant : BOUDANI Ali

Synthèse bibliographique

Aggregated Multicast – A Comparative study

Jun-Hong Cui, Jinkyu Kim, Dario Maggiorini, Khaled Boussetta, and Mario Gerla. Aggregated Multicast – A comparative study. *special issue of Cluster Computing: The Journal of Networks, Software and Applications, Baltzer Science Publisher, 2003*
http://www.cs.ucla.edu/NRL/hpi/AggMC/papers/aggmc_cc_journal.pdf

Master Pro. IR – IFISC décembre 2004

1. INTRODUCTION ET CONTEXTE.

L'intérêt du multicast est de diffuser des informations vers des groupes en optimisant les ressources du réseau (bande passante, routeurs). Le multicast d'IP est efficace au niveau de la transmission des données aux membres d'un groupe mais il souffre d'un problème d'adaptation au facteur d'échelle si l'on considère un très grand nombre de groupes actifs. En effet, cela demande aux routeurs de garder des informations pour chaque groupe. Pour résoudre ce problème, les auteurs proposent un schéma, appelé *Aggregated Multicast*. Tout d'abord ils nous présentent brièvement les modèles classiques du multicast puis expliquent leur schéma innovant. Ensuite, ils introduisent une métrique afin de mesurer les gains en terme de temps de traitement dans les routeurs, ainsi que la surcharge due aux différents arbres de routage. Enfin, ils testent les différents modèles par des simulations et concluent au vu des résultats.

Le routage multicast s'appuie sur la présence d'arbres de diffusion. Les stations membres du groupe forment les feuilles de l'arbre et sont les destinataires du paquet multicast. Les routeurs forment alors les noeuds de l'arbre. Ils appellent **noeuds terminaux**, les noeuds dont le trafic entre ou sort du domaine de routage, les autres noeuds sont appelés **noeuds de transit**. La gestion des arbres est tellement coûteuse qu'elle est peu déployée sur l'Internet.

Selon le type d'arbre distribué, nous classons les protocoles de routage en deux catégories:

- *source specific tree scheme* (exemple : DVMRP, PIM-DM et MOSPF)
- *shared tree scheme* (exemple : CBT,PIM-SM, BIDIR-PIM)

Avec *source specific tree scheme*, un arbre distribué est construit pour chaque source. Par contre, *shared tree scheme* construit un arbre pour chaque groupe et toutes les sources du groupe utilisent le même arbre pour acheminer les données aux récepteurs (en d'autres termes, plusieurs sources d'un même groupe partagent un unique arbre distribué).

2. AGGREGATED MULTICAST.

Dans leur schéma, *l'aggregated multicast*, plusieurs groupes multicast sont forcés de se partager un arbre distribué, qui est appelé **arbre agrégé**. En faisant cela, le nombre d'arbres dans le réseau est fortement réduit. Par conséquent, la tâche dédiée aux routeurs est diminuée : les routeurs ont seulement besoin de garder un état par arbre agrégé plutôt que par groupe.

En revanche, cette solution risque d'engendrer des pertes en terme de bande passante. En effet, des paquets pourront être acheminés à des noeuds non-membres. Il est évident qu'en fonction de la topologie du réseau et des groupes on obtient des arbres partagés plus ou moins performants. Lorsque l'arbre agrégé ne couvre pas totalement un groupe (ie certains noeuds terminaux pour un groupe ne sont pas des noeuds membres de l'arbre agrégé), on peut résoudre ce problème par un mécanisme de tunnelling. Un arbre agrégé peut être aussi bien *source specific tree* que *shared tree*.

3. SIMULATION ET MODELE.

Afin de mesurer les gains et pertes effectifs, ils ont dû introduire des métriques.

La première repose sur le **nombre d'arbres multicast**. Ceci reflète directement que plus il y a d'arbres multicast plus la charge des routeurs est importante.

La seconde est une indication sur le nombre **d'états d'acheminement** dans les noeuds de transit. En effet, cet état d'acheminement peut être réduit dans ces noeuds. Par contre, il ne peut

pas être réduit au niveau des noeuds terminaux car ces derniers ont besoin de maintenir des informations sur chaque groupe individuellement.

Ils ont mené leurs expériences avec SENSE (*Simulation Environment for Network System Evolution*) qui supporte aussi bien la source *specific tree*, le *shared tree*, et enfin *l'aggregated multicast* (aussi bien *shared* que *source specific*). Il faut dire qu'ils comparent les schémas et non les protocoles. Chaque schéma est implémenté avec une méthode centralisée (ie : présence d'une entité qui connaît toute la topologie du réseau).

Ils ont choisi un modèle pour lequel chaque noeud ne sera pas équivalent. De ce fait, chaque noeud se voit attribuer un poids qui correspond à la probabilité qu'il appartienne à un groupe multicast. Enfin, pour contrôler la simulation, ils ont choisi de commencer par fixer la taille du groupe, son espérance, et son taux d'arrivée. Ainsi, les données sont générées grâce à ces informations. Il est à noter que l'hypothèse a été faite de disposer de suffisamment de bande passante pour éviter de surcharger les liens.

Après avoir fixé tous ces paramètres, ils ont pu commencer les simulations en s'appuyant sur une topologie existante, le vBNS IP backbone qui est le réseau fédérateur internet américain à hautes performances réservé à la recherche et à l'enseignement. On peut décomposer ces tests en deux parties.

La première partie permet de comparer le *shared tree* à *l'aggregated multicast* (dans sa version *shared tree*). Ces tests montrent que *l'aggregated multicast* utilise beaucoup moins d'arbres (-50%), et de plus, la charge des routeurs est moindre (-40%). On observe également que plus l'on « perd » de bande passante et plus cette baisse est importante.

La deuxième partie a pour but de voir l'impact qu'ont la perte de bande passante et le tunnelling sur l'efficacité de *l'aggregated multicast*. Il est clair que plus on accepte de « perdre » de la bande passante et plus on utilise le mécanisme de tunnelling et plus l'agrégation est efficace.

Ils ont obtenu des résultats similaires dans les autres cas comme l'*aggregated* (version *source specific tree*) face au *source specific tree*. En fait, dans tous les cas, l'*aggregated multicast* se révèle plus performant en terme de réduction du nombre d'arbres et des états de transit.

4. CONCLUSION.

Au vu de tous ces résultats, on peut dire que le bénéfice de *l'aggregated multicast* repose principalement sur deux points :

- Réduction du travail des routeurs par diminution du nombre d'arbres,
- Réduction des états à maintenir au niveau des noeuds de transit.

En revanche, le prix à payer est une perte de bande passante, ainsi que le coût engendré par le mécanisme de tunnelling. C'est même plus fort que cela, car plus on accepte de « perdre » de la bande passante et plus on utilise le mécanisme de tunnelling et plus l'agrégation devient efficace.

Pour résoudre le problème de mise à l'échelle, les auteurs proposent donc une méthode innovante : *l'aggregated multicast* dont l'idée clé est de forcer les groupes à partager un même arbre. Ils ont même été plus loin car ils ont proposé une métrique pour mesurer son efficacité, et vérifié expérimentalement.