

## **SETI@Home et le calcul distribué volontaire au service de la science**

Jimmy Broché, Aureliano Naviliat, Marie Rodriguez, Loïc Vaes  
ULB - Département d'Informatique

L'informatique est une science au service des autres disciplines scientifiques; elle est devenue indispensable dans de nombreux domaines de recherche actuels, depuis l'analyse des données produites par les accélérateurs de particules jusqu'aux simulations qui permettent de prédire les changements climatiques.

Nombre de ces applications nécessitent de très grandes capacités de traitement de l'information. Pour y faire face, les institutions concernées doivent souvent faire le choix entre l'acquisition d'un supercalculateur, c'est-à-dire un seul ordinateur très puissant, ou bien d'un groupe d'ordinateurs moins puissants qui collaborent et se partagent la charge de travail. Cette deuxième solution, le calcul distribué, est souvent plus souple et plus intéressante financièrement.

Certains projets ont néanmoins développé une autre approche, le recours à la participation de volontaires. Des particuliers répartis dans le monde entier mettent à disposition des chercheurs la capacité de calcul de leurs ordinateurs personnels, typiquement pendant les périodes d'inactivité de leurs processeurs, quand ils sont en veille par exemple. Dans ce modèle, l'organisme de recherche distribue à ses participants de petites unités d'informations à analyser, celles-ci sont traitées par les ordinateurs des volontaires, et les résultats sont renvoyés à l'équipe scientifique.

Bien que la puissance individuelle des machines utilisées soit relativement faible, le nombre de participants permet de rivaliser avec les supercalculateurs les plus puissants, à un coût modéré pour l'équipe de recherche. En fait, un de ces projets regroupe plusieurs millions de personnes, et sa puissance totale de calcul est deux fois plus grande que celle du plus puissant supercalculateur au monde, lequel a coûté plus de soixante millions d'euros.

SETI@home est un de ces projets de calcul distribué volontaire, et le pionnier de ce système. SETI, *Search for Extraterrestrial Intelligence*, a pour mission de chercher l'existence de civilisations extraterrestres en tentant de détecter des signaux radios émis depuis l'espace. Des télescopes balaient la voûte céleste, et produisent une très grande quantité de données qui sont segmentées et réparties entre les volontaires pour traitement. Les ordinateurs des participants analysent les informations reçues, et informent l'équipe de SETI des résultats obtenus. Si jusqu'ici aucun signal extraterrestre n'a été détecté, la recherche continue.

Ce type d'organisation présente néanmoins quelques inconvénients, dont la question de la fiabilité des données obtenues. Si les calculs sont effectués par des volontaires et non sur des ordinateurs qui appartiennent à l'institution en question, il est impossible de contrôler le traitement des données, et difficile d'être certain des résultats reçus. Des résultats incorrects peuvent être causés par exemple par des erreurs de calcul au niveau du processeur, par des bogues dans les logiciels utilisés, ou même potentiellement par des erreurs introduites volontairement.

Différents mécanismes ont heureusement été développés pour résoudre ce problème. Un de ceux-ci est la duplication des traitements, c'est-à-dire l'envoi à plusieurs volontaires des mêmes données, et la comparaison des résultats obtenus, pour essayer de déterminer par vote à la majorité quelles sont les réponses correctes. C'est cette solution qui est utilisée par SETI@home et étudiée dans le cadre de cette présentation. D'autres techniques comprennent notamment des tests ou quizz, dans lesquels les participants sont périodiquement testés par l'envoi de données dont l'équipe scientifique connaît en fait déjà les résultats.