



Michael Stonebraker, ACM Turing Award pour ses contributions au domaine de la gestion de données

Patrick Valduriez¹

Michael Stonebraker, chercheur au Massachusetts Institute of Technology (USA), vient de remporter le prestigieux ACM Turing Award, souvent considéré comme « le prix Nobel de l'informatique ». Dans son annonce du 25 mars 2015, l'ACM précise que Stonebraker « a inventé de nombreux concepts qui sont utilisés dans presque tous les systèmes de bases de données modernes... ». Cette reconnaissance au plus haut niveau international me donne l'occasion de donner un éclairage sur la place singulière du domaine de la gestion de données dans la recherche en informatique.

La gestion de données se préoccupe du stockage, de l'organisation, de la recherche et de la manipulation de données de toutes sortes, souvent très grandes et complexes. C'est aujourd'hui un domaine majeur de l'informatique, avec à la fois une grande communauté de recherche et une industrie forte. En particulier, le transfert continu de la recherche vers l'industrie a conduit au développement des systèmes de gestion de bases de données (SGBD), au cœur de tout système d'information moderne.

Le principe fondateur de la gestion de données est l'*indépendance des données*, qui permet de travailler avec les données à un niveau conceptuel, tout en ignorant les détails d'implémentation. Le modèle relationnel, en s'appuyant sur une théorie

1. Directeur de Recherche Inria, ACM Fellow 2012.

simple et solide (ensembles, logique du 1^{er} ordre) a révolutionné la façon de concevoir les SGBD. Il a aussi favorisé le nécessaire rapprochement entre théorie et pratique des bases de données, qui s'est concrétisé par la fusion de grandes conférences internationales comme ACM SIGMOD/PODS et EDBT/ICDT.

L'innovation majeure des SGBD relationnels a été de permettre la manipulation de données avec des langages de requêtes déclaratifs, intégrant des concepts puissants comme les transactions ou les règles d'intégrité, et facilitant l'optimisation automatique. Arrivés sur le marché dans les années 1980, les SGBD relationnels ont remarquablement réussi le test du temps, par l'ajout régulier de nouvelles fonctionnalités (p. ex. sécurité, auto-réglage), de nouveaux types de données (p. ex. objet, XML ou JSON) et en s'adaptant à toutes sortes de plateformes, depuis les appareils mobiles (p. ex. smartphones) jusqu'aux très grands clusters dans des environnements distribués.

Aujourd'hui, avec les nouveaux défis du *big data* et du *cloud*, la gestion de données doit être réinventée, tant les besoins des utilisateurs et des applications sont divers et ne peuvent plus s'accommoder de l'aspect « taille unique » des SGBD relationnels. La recherche en gestion de données devient alors pluridisciplinaire, associant notamment chercheurs et grands producteurs de données pour mieux étudier leurs données (*big data analytics*). En France, ces défis sont au cœur d'initiatives pluridisciplinaires récentes comme le défi CNRS Mastodons (grandes masses de données scientifiques) et le GdR MaDICS (Masses de Données, Informations et Connaissances en Sciences).

En 40 ans de carrière, Stonebraker a profondément marqué le domaine des SGBD, depuis le relationnel jusqu'au *big data*. Sa récompense s'ajoute à celles de trois autres *ACM Turing Awards* du domaine : Charles Bachman (1973) pour ses contributions aux SGBD navigationnels, Edgar Frank Codd (1981) pour l'invention du modèle relationnel et James Gray (1998) pour ses contributions aux SGBD et au transactionnel.

Stonebraker a d'abord été un pionnier dans la conception de SGBD relationnels, en développant deux des projets de recherche les plus influents au monde, Ingres et Postgres (relationnel et objet), à l'université de Berkeley. Puis il a conçu de nouveaux SGBD comme Mariposa (SGBD distribué), Streambase (SGBD pour flux de données), Vertica (SGBD orienté-colonne pour le décisionnel) et SciDB (SGBD orienté-tableau pour les applications scientifiques). Il a été aussi un entrepreneur exceptionnel, en créant neuf startups autour de ses projets dont Relational Technology Inc. (Ingres), devenue aujourd'hui Ingres Corp., Illustra (à partir de Postgres) acquise par Informix, Cohera (à partir de Mariposa) acquise par PeopleSoft, Streambase acquise par Tibco et Vertica acquise par HP. Enfin, il a créé deux grands projets en logiciel libre : PostgreSQL et SciDB. Aujourd'hui, Stonebraker poursuit ses travaux au MIT autour des systèmes NoSQL.