

IBM investit massivement dans la technologie Big Data Spark

Apache Spark est un projet en pleine croissance dans le monde du **Big Data**. Cette offre est capable de fonctionner sur un cluster de serveurs et permet de traiter les données **jusqu'à 100 fois plus rapidement que Hadoop**, du fait de l'utilisation de techniques *In-Memory*.

IBM n'a pas manqué de remarquer cette solution Open Source, que la firme juge comme étant potentiellement « *le plus important nouveau projet Open Source de la décennie* » dans le secteur du traitement de données.

Big Blue compte donc mettre le paquet sur Spark. Cette offre sera intégrée à ses solutions dédiées aux mondes **de l'analytique et du commerce**. Elle sera également proposée sur le Watson Health Cloud et déclinée en offre à la demande **sur le PaaS Bluemix**.

SystemML sera placé sous licence Open Source

Dans le même temps, **plus de 3 500 développeurs et chercheurs d'IBM** se pencheront sur des projets relatifs à Spark. 12 laboratoires seront investis dans ces travaux, dont un – situé à San Francisco – dédié spécifiquement à Spark.

La technologie de Machine Learning d'IBM, **SystemML**, sera également offerte à la communauté Spark sous forme Open Source. Enfin, diverses opérations de formation seront proposées au public, avec comme objectif de former **plus d'un million de data scientists et data engineers** à Spark, au travers de cours en ligne (MOOC).

À lire aussi :

[Le gouvernement britannique s'associe avec IBM sur le Big Data](#)

[Des rachats pour IBM et Cisco : ruée vers OpenStack](#)

[Apache vient booster Hadoop avec Spark 1.0](#)

Crédit photo : © Lassedesignen – Shutterstock