

# Amazon Web Services intègre Spark à son offre Cloud Hadoop

Le projet Open Source **Spark** d'Apache a décidément le vent en poupe. Après IBM, qui a annoncé un investissement d'envergure dans cette technologie (voir l'article « [IBM investit massivement dans la technologie Big Data Spark](#) »), c'est aujourd'hui au tour **d'Amazon** d'adopter cette offre.

La firme joue la carte du couple **Hadoop + Spark**, en proposant Spark par-dessus son offre Big Data **Amazon Elastic MapReduce** (EMR). Le support de Spark est natif dans EMR, et donc directement accessible (les utilisateurs devaient auparavant l'installer manuellement).

Spark permettra de disposer d'un surcroît de puissance, via l'utilisation **de techniques *in-memory*** de traitement de données, permettant de stocker l'ensemble du processus en mémoire : jeu de données initial, résultats intermédiaires et résultat final.

## Prêt pour le machine learning

*« Spark supporte nativement Scala, Python et Java. Il inclut des bibliothèques pour le SQL, des algorithmes pour le machine learning, le traitement orienté graphe et le traitement des flux. Avec de nombreuses options de développement intégrées, il peut être plus facile de créer et maintenir des applications pour Spark que de travailler avec les diverses abstractions proposées autour de l'API Hadoop MapReduce, »* explique **Jon Fritz**, senior product manager chez Amazon, [sur le blog d'AWS](#).

### À lire aussi :

[DaaS, Machine Learning, conteneurs, stockage : Amazon booste son Cloud](#)

[Résultats : Amazon confirme sa place de leader sur le Cloud public](#)

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