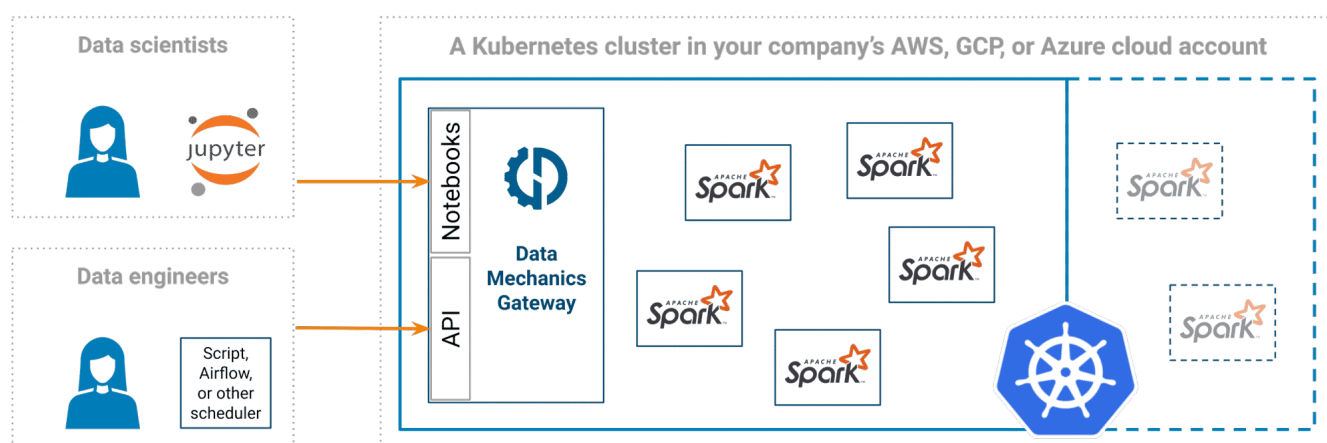


# Spark sur Kubernetes : Data Mechanics tape dans l'œil de NetApp

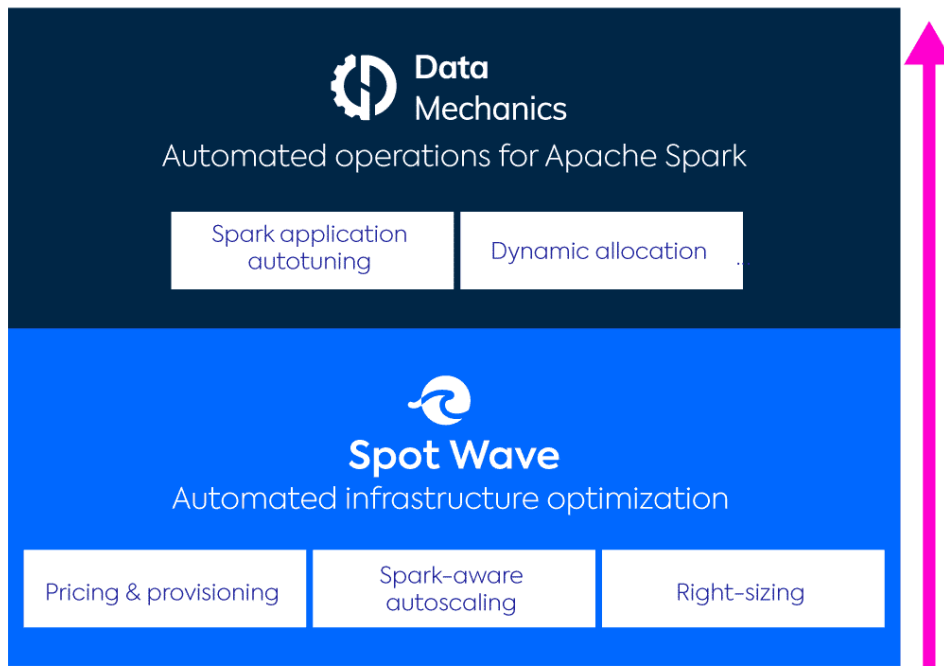
Spot et Data Mechanics, une alliance idéale ? Les deux entreprises vont en tout cas fusionner – et leurs technologies avec. La première, à l'origine d'outils de gestion d'infrastructure cloud, était [tombée](#) dans l'escarcelle de NetApp il y a un an. La seconde vient d'annoncer qu'elle allait [suivre](#) le même chemin. Avec, dans ses bagages, une version managée de Spark sur Kubernetes.

Cette offre concurrence notamment celle de Databricks... dont provient Jean-Yves Stéphan, le principal dirigeant de Data Mechanics. Elle se positionne aussi face aux solutions des *hyperscalers* : Dataproc (Google), EMR (Amazon) et HDInsight (Microsoft). C'est d'ailleurs sur leurs plates-formes qu'elle se déploie – il n'existe pas encore de version *on-prem*.



Spark sur Kubernetes est considéré comme stable depuis la version 3.1 du *framework*, [sortie](#) il y a quelques semaines. En plus de l'aspect managé (création des clusters, *autoscaling* avec gestion des instances *spot*, monitoring, sécurité...), Data Mechanics ajoute :

- La prise en charge native des conteneurs Docker
- La collecte et la conservation des *logs* et des métriques
- Un mode autopilote qui s'appuie sur l'historique des pipelines Spark pour ajuster les configurations
- Diverses intégrations ; par exemple [avec](#) Jupyter et Airflow pour soumettre des tâches



La start-up a par ailleurs développé une UI alternative pour Spark. Elle en a fait un [projet open source](#), sous la marque Delight. Sa tarification se fonde non pas sur le temps de fonctionnement, mais sur le temps de calcul effectif.

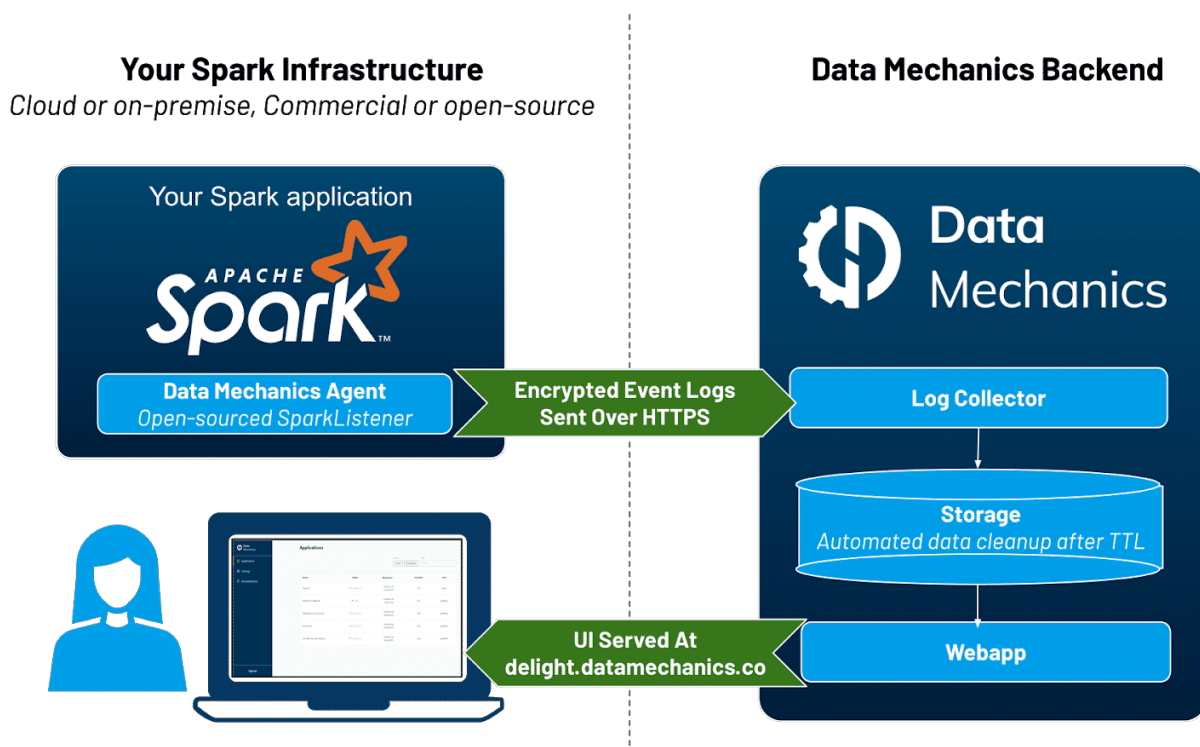


Illustration principale © Markus Spiske – Unsplash