



COMMUNIQUÉ DE PRESSE

Twitter : Teradata : Think Big développe ses offres de data lakes avec Apache Spark #teradata

Think Big, une société Teradata, étend les fonctionnalités de création de data lakes avec Apache Spark

Les défis liés au déploiement de Spark engendrent une demande croissante des services big data de Teradata dans le monde entier

Antony, France, le 14 avril 2016 — [Teradata](#) (NYSE : [IDC](#)), le spécialiste de l'[analyse big data](#) et des [applications marketing](#), annonce que Think Big, cabinet-conseil Teradata proposant dans le monde entier son expertise de leader dans le déploiement d'Apache Spark™ et d'autres technologies big data, développe ses offres de data lakes et de services gérés grâce à Apache Spark. Spark est une plate-forme informatique open source à cluster utilisée pour les recommandations produit, l'analyse prédictive, l'analyse des données de capteurs, l'analyse graphique et bien plus.

Les clients peuvent actuellement utiliser un data lake avec Apache Spark dans le cloud, sur des environnements Hadoop génériques « à base de matériel standard », ou avec [Hadoop Appliance de Teradata](#), la plus puissante plate-forme d'entreprise, prête à l'emploi, préconfigurée et optimisée pour exécuter des charges de travail big data professionnelles.

Alors que l'intérêt pour Spark continue de croître, de nombreuses entreprises peinent à suivre le rythme rapide des changements et la fréquence des versions de la plate-forme open source. Think Big a intégré avec succès Spark dans ses systèmes pour réaliser des data lakes et des applications analytiques de qualité professionnelle.

« De nombreuses entreprises expérimentent Apache Spark dans l'espoir de tirer parti de ses points forts en matière de streaming de données, de requêtes et d'analyse – souvent en combinaison avec un data lake », constate Philip Russom, directeur de la recherche en gestion de données chez The Data Warehousing Institute (TDWI). « Cependant les utilisateurs réalisent rapidement que Spark n'est pas facile à utiliser et que les data lakes nécessitent davantage de

planification et de conception qu'ils ne le pensaient. Les utilisateurs se retrouvant dans cette situation doivent faire appel à de l'aide extérieure auprès de consultants ou de fournisseurs de services gérés ayant acquis de l'expérience avec Apache Spark et les data lakes dans le cadre de projets réussis avec divers clients. Think Big possède précisément cette expérience ».

Think Big élabore des packages de services reproductibles pour le déploiement de Spark, y compris pour l'ajout de Spark en guise de moteur d'exécution pour ses offres de data lake et de services gérés. Par l'intermédiaire de sa branche formation – Think Big Academy – le cabinet lance également une série de nouvelles offres de formation Spark pour les entreprises clientes. Encadrés par des instructeurs expérimentés, ces cours contribuent à former les gestionnaires, les développeurs et les administrateurs à l'utilisation de Spark et de ses différents modules, notamment l'apprentissage machine, les graphiques, le streaming et les requêtes.

Par ailleurs, l'équipe d'ingénierie des données de Think Big va diffuser en open source les routines pour le clustering [K-Modes](#) distribué avec l'API (interface de programmation d'applications) Python de Spark. Ces routines améliorent le clustering des données catégorielles pour la segmentation de la clientèle et l'analyse de la perte de clients. Ce code sera disponible avec d'autres travaux open source de Think Big sur la page [GitHub](#) de Think Big.

« Notre cabinet-conseil Think Big se développe rapidement, aussi bien en Amérique qu'en Europe et en Chine, en raison de l'explosion de la demande d'expertise, d'expérience et de méthodes pour aider les entreprises à se doter d'un data lake en utilisant Spark et Hadoop de façon adéquate et du premier coup », déclare Ron Bodkin, le président de Think Big. « Le déploiement de Spark doit s'inscrire dans une stratégie d'information et d'analyse. Nous savons par expérience quels cas d'utilisation sont pertinents, quelles sont les bonnes questions et où se trouvent les écueils à éviter lors du déploiement. Nous maîtrisons aussi bien les attentes des utilisateurs professionnels que les exigences technologiques. Nous pouvons contribuer à générer une valeur ajoutée concrète et nos clients Spark le font déjà dans des domaines allant de la personnalisation client omnicanal à la détection d'erreur en temps réel dans la fabrication de haute technologie. »

Bien avant que le big data ne devienne la tendance à la mode, Think Big était déjà la première et la principale société de services exclusivement dédiée au big data et mettant en œuvre des solutions analytiques basées sur des technologies émergentes. Aujourd'hui, Think Big fournit des services gérés pour Hadoop dans le domaine de la prise en charge de plates-formes et d'applications, à l'aide de processus bien définis, d'outils solides et de spécialistes expérimentés

en big data, afin de gérer, de contrôler et d'entretenir la plate-forme Hadoop à un prix abordable. En débutant chaque nouvelle mission par un processus de transition soigneusement testé, Think Big évalue et optimise les équipes d'assistance à la production, de développement et d'entretien d'un client pour un déploiement le plus efficace possible.

Liens d'actualités connexes (en anglais)

- Services de mise en œuvre Think Big SPARK : pour plus de détails [visitez le site web de Think Big](#)
- Teradata arrive en position de leader dans l'édition 2016 du Gartner Magic Quadrant for Data Warehouse and Data Management Solutions for Analytics – [Consultez le nouveau rapport ici](#)

À propos de Teradata

Teradata (NYSE : TDC) aide les entreprises à tirer profit de leurs données plus qu'aucune autre entreprise. La gamme de solutions leader de Teradata en matière d'analyse big data, d'applications marketing intégrées et de services peut aider les entreprises à tirer de leurs données un avantage concurrentiel durable. Consultez le site teradata.com.

###

Faites connaissance avec Teradata :



Teradata et le logo de Teradata sont des marques commerciales ou des marques déposées de Teradata Corporation et/ou de ses filiales, aux États-Unis et dans le reste du monde.

Contacts Presse:

Charles-Yves Baudet
Teradata Corporation Public Relations
+ 33 1 64 86 76 14
charles-yves.baudet@teradata.com

Victoria Bareille
Réflexion Publique
+ 33 (0) 1 55 21 01 55
victoriab@reflexion-publique.com