

Moteur d'analyse pour les données distribuées

Alimenté par Starburst Enterprise

Le chemin le plus court de la dispersion des données vers de meilleures décisions

Le plus souvent, les organisations utilisent plusieurs bases de données et systèmes de stockage pour entreposer leurs données. L'un des plus grands défis, que vous mettiez en œuvre une nouvelle stratégie de données ou que vous mainteniez une stratégie existante, est la dispersion des données. Les données sont partout! Comment pouvez-vous regrouper vos données et les rassembler en un seul endroit pour prendre des décisions commerciales plus rapides, meilleures et basées sur les données ?

L'évolution de la modernisation des solutions Big Data a tenté de relever ce défi, alors que le volume des données augmente tous les jours de manière exponentielle et que les écosystèmes de données deviennent plus complexes.

La première vague d'évolution a commencé avec les entrepôts de données où l'extraction, la transformation et le "chargement" (ETL) des données dans le data warehouse pour un usage spécifique sont devenus populaires. Cela fonctionne si vous pouvez imaginer, extraire, transformer et charger les bonnes données dans le data warehouse. Mais il n'est tout simplement pas réaliste de se poser des questions encore inconnues ou de connaître les sources de données qui seront nécessaires pour des informations qui seront disponibles à l'avenir. À mesure que de nouvelles questions se posent, les ingénieurs de données s'adaptent rapidement afin de pouvoir utiliser des données provenant de nouvelles sources et exécuter le processus ETL. Cela retarde encore plus les informations, car le consommateur de données doit attendre l'achèvement de l'ETL avant de procéder à son analyse de données.

La vague suivante a été l'introduction du data lake. Le data lake a éliminé le processus ETL, car il a fourni un data store où vous pouvez inclure à la fois des données structurées et non structurées à grande échelle. Malheureusement, les data lakes sont sollicités par le volume de données que vous y déchargez pour une utilisation future indéfinie. Cette option devient coûteuse à mesure que le volume de données brutes augmente. De plus, le mouvement des données peut entraîner des situations dans lesquelles les données sont obsolètes.

Le data lakehouse semble être une option pour les problèmes de gestion des données, car il combine des avantages du data warehouse et du data lake. Il prend en charge à la fois les données traitées et les données brutes. Il a également de faibles coûts de stockage dans le cloud tout en fournissant des capacités de requête avancées. Les données traitées peuvent être organisées en graphiques et en tableaux tout en permettant de stocker des données brutes ou d'autres formes de données plus récentes dans un emplacement central. Mais pouvons-nous vraiment obtenir tous les avantages analytiques du data warehouse et les avantages de stockage du data lake ? Ou sera-t-il bientôt une fois de plus évident que la « solution » à la gestion des données soit en dessous des normes promises par le mythe d'une source unique de vérité ? L'avenir nous le dira.

Problèmes actuels des données dans l'entreprise:

- Absence d'analyse holistique des données pour une prise de décision plus rapide et de meilleure qualité
- Augmentation des coûts et de la complexité des données
- Mauvaise performance avec de très grands volumes de données
- Respect de la sécurité et de la gouvernance de l'entreprise

Et si vous pouviez utiliser les systèmes de stockage que vous possédez actuellement et les faire fonctionner pour vous, plutôt que contre vous ? Et si on pouvait laisser les données là où elles sont ? Ou mieux encore, extraire, transformer et charger certaines données une dernière fois dans un lieu de stockage moins cher ? Et si, au lieu d'une source unique de vérité, nous avions un point d'accès unique à toutes les données ?

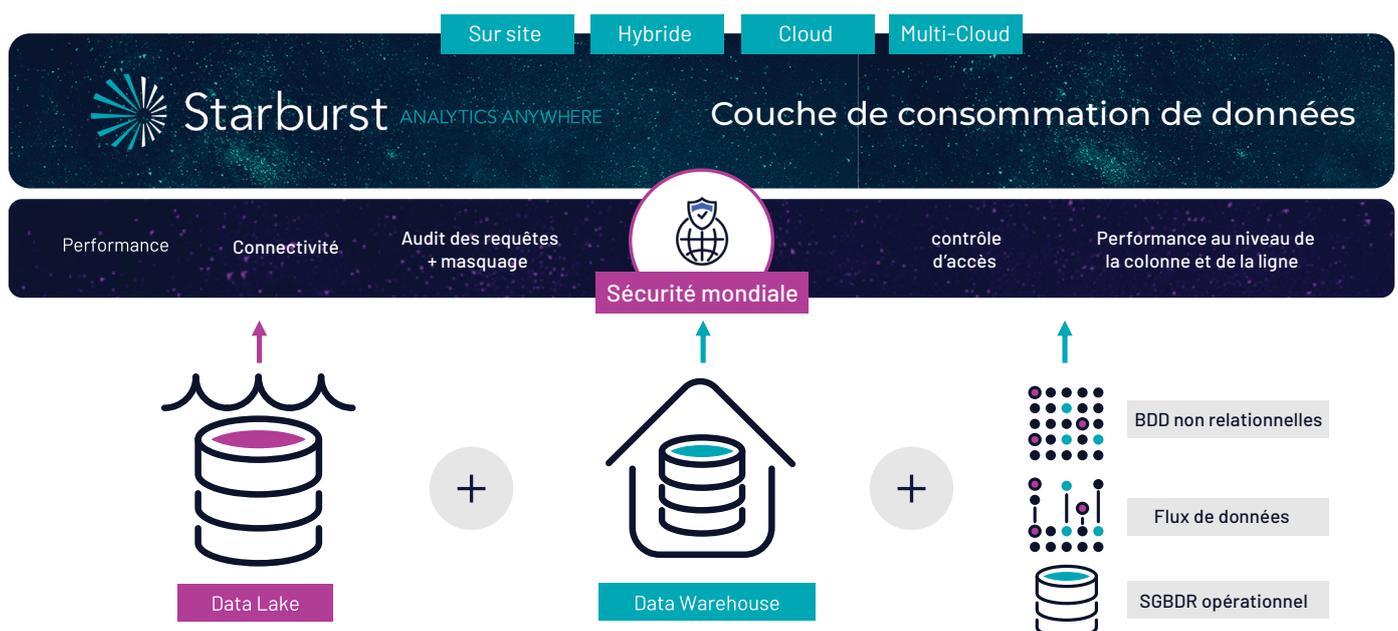
Starburst comme moteur d'analyse de vos données distribuées

Ne cherchez pas plus loin. Starburst offre aux plus grandes organisations du monde des capacités importantes qui résolvent ces challenges, notamment:

-  **Fédération:** rassemble facilement plusieurs sources de données, qu'elles soient on-premise ou dans n'importe quel cloud, dans un data lake, un data warehouse ou un lakehouse
-  **Pas de mouvement/duplication des données:** réduit ou élimine le besoin d'ETL, diminue le coût total de possession (TCO) et garantit la qualité de vos données tout en effectuant des requêtes ad hoc
-  **Optionnalité:** le verrouillage des données n'est pas une option, où que se trouvent vos données. Nous vous aidons à débloquer l'accès à toutes vos sources de données pour accélérer le temps d'obtention de vos informations
-  **Moteur SQL rapide pour le data lake:** accélère les temps d'accès aux informations sur votre data lake grâce à notre moteur de requête SQL ANSI
-  **Échelle et concordance:** mise à l'échelle avec votre volume de données croissant afin que vous puissiez réduire considérablement l'utilisation des ressources et fournir une concordance adéquate

Comment ça fonctionne

Starburst résout le problème de l'intersection de volumes de données sans cesse croissants et de l'augmentation du nombre de consommateurs de données. Notre point de vue est que l'approche historique consistant à déplacer constamment les données dans une méthode de source unique de vérité n'est plus possible : il y a trop de données et trop de personnes essayant de poser de requêter les données. Nous aidons nos clients à dé-prioriser les mouvements de données tout en donnant à leurs équipes un accès SQL cohérent, facile et rapide à toutes leurs données. C'est ce que nous appelons la couche de consommation de données.



La couche de consommation de données

Starburst Enterprise est une couche de consommation des données qui déverrouille les données en silo en fournissant un accès rapide à toute source de données, via SQL, le langage que votre équipe connaît et qui peut être appliqué dans l'outil d'analyse que vous utilisez déjà. Avec des fonctionnalités de sécurité avancées, cette couche de consommation de données fournit un point d'accès unique sécurisé à toutes vos données pour tous vos consommateurs de données.

Les connecteurs Starburst Enterprise haute performance permettent des requêtes à partir d'une ou de plusieurs sources, on-premise ou dans le cloud, et Starburst Enterprise Presto tout est régi par des fonctionnalités de sécurité mondiale de qualité professionnelle avec un contrôle d'accès précis.

Sécurité de qualité professionnelle

L'architecture légère et autonome de Starburst Enterprise simplifie l'installation, la sécurisation, la maintenance et la mise à l'échelle. Puisqu'il n'y a pas de stockage de données et qu'il peut être installé dans n'importe quel emplacement, y compris dans le cloud ou sur site, la sécurité est simple à assurer et à appliquer. Starburst offre des fonctionnalités de sécurité de qualité professionnelle à l'échelle mondiale conçues pour répondre aux normes des plus grandes organisations au monde.

Les améliorations en matière de sécurité comprennent :

- Contrôle d'accès précis au niveau du connecteur, de la table, de la colonne et de la ligne
- Le chiffrement des données de bout en bout, l'authentification des utilisateurs et d'autres pratiques exemplaires de sécurité d'entreprise protègent davantage vos clusters et les données interrogées
- Intégration de Kerberos et de LDAP, masquage des données et audit des requêtes

Écouter ce que nos clients ont à dire



Starburst nous a vraiment permis d'accélérer les temps d'obtention des informations, d'améliorer nos taux de conversion et de nous permettre une modélisation plus robuste. Tous ces facteurs nous ont aidés à obtenir de meilleurs résultats commerciaux.

Mitchell Posluns
Spécialiste sénior de l'analyse de données, Assurance



Il est tellement plus facile de fournir aux utilisateurs un seul point d'accès. Ils peuvent utiliser les mêmes outils qu'ils connaissent bien, mais tout se fait plus vite.

Mike Prior
Ingénieur E/S principal Optum

En savoir plus

- Visitez notre [Starburst Enterprise](#) et téléchargez un essai gratuit
- Passez en revue le [mémoire de solution technique Starburst](#) et le [guide de sécurité Starburst Enterprise](#)
- Besoin de plus d'informations? [Contactez-nous!](#)

