

IBM Db2 Big SQL: proporciona acceso SQL a su big data

Big data es un gran imperativo para el negocio, y los analistas coinciden en que SQL es la clave para facilitar el desbloqueo de valor desde Apache Hadoop. Hadoop introduce nuevas tecnologías, incluyendo motores de analítica, lenguajes de script e interfaces de programación de aplicaciones (API). Con la expansión de nuevas tecnologías en el ecosistema de Hadoop, es muy importante elegir una plataforma Hadoop que esté en consonancia con los estándares abiertos, como la Open Data Platform initiative (ODPi). Hortonworks Data Platform (HDP) cumple con la ODPi, y Big SQL está diseñado para ejecutarse en HDP, reforzando el valor de ODPi.

Con HDP como base, Big SQL permite a los usuarios consultar datos de Hive y HBase utilizando SQL en cumplimiento con ANSI. Si bien Hadoop es muy escalable, la arquitectura de procesamiento en paralelo masivo (MPP) y el

optimizador avanzado basado en costes de Big SQL pueden ejecutar consultas de forma más inteligente, no más complicada, dando soporte a más usuarios simultáneos y SQL más complejo con menos hardware, en comparación con otras soluciones SQL para Hadoop. Big SQL también forma parte de las soluciones de analítica de big data de IBM, tales como IBM Big Replicate, IBM BigIntegrate e IBM BigQuality.

Big SQL es a su vez una potente plataforma para la descarga y unificación de almacenes de datos, básico para muchos usuarios de Hadoop. BigSQL es la primera y única solución SQL-on-Hadoop que comprende la sintaxis SQL de uso más frecuente de otros proveedores y productos, como Oracle, IBM Db2 e IBM Netezza. Y, si los datos no se pueden mover a Hadoop, Big SQL proporciona acceso federado a orígenes de sistemas de gestión de bases de datos relacionales (RDBMS) externos a Hadoop con la tecnología IBM Fluid Query.

Principales características de IBM Big SQL

- **Rendimiento SQL-on-Hadoop superior** utilizando la tecnología de impulso elástico
- **Capacidad para comprender los dialectos SQL de otros proveedores y productos** para Oracle, IBM® Db2® e IBM Netezza®, lo que facilita y acelera la descarga de datos antiguos de almacenes de datos empresariales existentes o data marts para liberar capacidad, conservando la mayor parte de SQL de estas plataformas
- **Un motor híbrido para Hadoop** que utiliza Hive, HBase y Apache Spark simultáneamente para ofrecer las mejores funcionalidades de analítica
- **Mayor integración** con Spark que otras tecnologías SQL-on-Hadoop, habilitando nuevos casos de uso
- **Seguridad avanzada de filas y columnas:** la integración de Big SQL y Spark mejora la seguridad de la "compra de datos" ya que permite enmascarar atributos sensibles de forma predeterminada sin puertas traseras a los datos, habilitando el acceso de autoservicio a los datos de forma segura y controlada

Hadoop dispone de múltiples motores de analítica de última generación para resolver problemas de big data. Big SQL ofrece un rendimiento SQL-on-Hadoop superior con la tecnología de impulso elástico. Big SQL ofrece también integración completa de Spark 2.1, lo que habilita nuevos casos de uso. ¿Necesita exploraciones, inserciones, actualizaciones y supresiones de alto rendimiento? ¿Necesita machine learning o analítica de gráficos con Spark, con un único modelo de seguridad? Big SQL es un motor SQL para Hadoop que aplica simultáneamente Hive, HBase y Spark mediante una única conexión de base de datos—incluso una sola consulta. Por ello, Big SQL es a su vez un eficiente motor híbrido.

¿La mejor parte? Todos los datos pertenecen a Hadoop. Las tablas de Big SQL son tablas de Hives, tablas de HBASE o conjuntos de datos distribuidos resilientes (RDDs) de Spark y se integran con Hive Metastore. Ahorre costes descargando los datos en Hadoop y aplique la analítica adaptada de última generación con el optimizador SQL para Hadoop.

Para obtener más información sobre IBM Big SQL, póngase en contacto con su representante de IBM o IBM Business Partner, o bien visite: ibm.biz/ibmbigsq