

Amazon Redshift, la revolución en Data Warehousing

Amazon sigue expandiendo su oferta de servicios basados en cloud computing a través de la línea [AWS \(Amazon Web Services\)](#). En esta ocasión ha presentado una nueva oferta para el **almacenamiento de grandes volúmenes de datos relacionales** denominada [Amazon Redshift](#).

Básicamente se trata de un **Data Warehouse** optimizado para albergar bases de datos relacionales provenientes de múltiples fuentes y sistemas y ofrecer un alto rendimiento en procesos de análisis e informes. En este tipo de infraestructuras se recopilan los datos de transacciones financieras, de ventas, de cadenas de suministro... y se unifican para que los responsables de las compañías puedan analizarlos y tomar las decisiones correctas para los negocios.

Tal y como explica la compañía, el mayor problema que tienen estos sistemas es que no es fácil ni barato mantenerlos, escalarlos y ejecutarlos ya que los modelos tradicionales cuentan con múltiples recursos de hardware, software, almacenamiento y redes. Partiendo de esa base ha desarrollado Redshift, que se base en los modelos que ha estado implementado en sus Servicios Web basados en cloud computing.

“Haciendo cálculos encontramos que una oferta media viene a costar entre 19.000 y 25.000 dólares por TeraByte por año. A través de Amazon Redshift los clientes pagarán **menos de 1.000 dólares por TeraByte por año**, siempre con los mismos niveles de servicio y escalabilidad pero eliminando los quebraderos de cabeza asociados con la administración y la monitorización cuando se gestiona un Data Warehouse propio”.

De ser correcta esta información, estaríamos hablando de una oferta realmente atractiva para aquellas empresas que necesitan manejar datos del orden de PetaBytes.



El proceso de configuración de Amazon Redshift (clic para ampliar)

En el comunicado hecho público desde el [blog de Amazon Web Services](#) explican que el proceso de puesta en marcha no puede ser más sencillo:

1. Aprovisionar un cluster a través de la consola de gestión AWS
2. Instalar los controladores (JDBC y ODBC)
3. Obtener un cliente SQL
4. Cargar los datos

En el lado de la seguridad e integridad de los datos, Amazon propone el almacenamiento en dos instancias hardware exactas, de tal forma que si una falla (los discos están continuamente monitorizados), se cambia a la otra automáticamente.