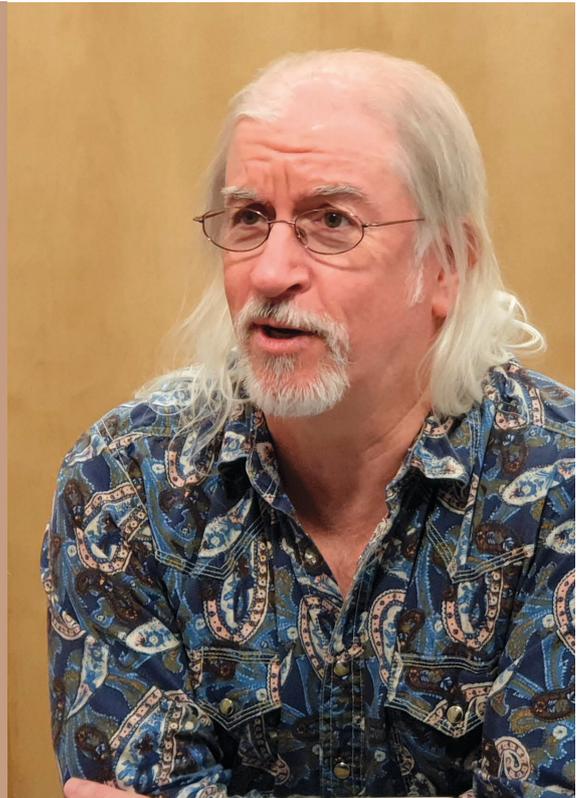


“La nube es el futuro de la gestión de los datos”

En Data Center Market hemos tenido la oportunidad de hablar con Kent Graziano, chief technical evangelist y strategic advisor en Snowflake, una compañía dedicada al mundo del data warehouse como servicio que apenas lleva unos meses instalada en España, pero que llega con ganas de crecer y cambiar la forma en la que los datos son gestionados y analizados.



Lucía Bonilla

✉ lucia.bonilla@bps.com.es
🐦 @DataCenterBPS
🌐 www.datacentermarket.es

Todos somos conscientes de la explosión del dato, no obstante, las compañías siguen teniendo muchos problemas para gestionarlos y obtener valor. ¿Dónde cree que está el problema?

Es una combinación de varios factores. Es verdad que el volumen de datos se ha incrementado drásticamente, y ya estamos hablando de cientos de Zetabytes, y que además provienen de múltiples fuentes, como dispositivos móviles, IoT y sensores conectados, coches conectados, incluso electrodomésticos. Estamos tan abrumados por los datos que hace diez años el problema que se planteaba era buscar la forma de almacenarlos, y cómo hacerlo más eficiente desde el punto de vista de costes. Hoy eso ya no es un problema: la cuestión está en qué hacer con esos datos. Y la tecnología hasta hace poco no era capaz de aportar soluciones a la hora de procesarlos y gestionarlos adecuadamente.

Ahí es donde entra Snowflake: sus fundadores estaban buscando oportunidades en la nube y en aquellas tecnologías que permitieran hacer algo

diferente para poder gestionar los datos pero que a la vez facilitara incorporar el mundo tradicional de bases de datos y data warehouse para hacer una integración real y conseguir que el almacenamiento fuera menos costoso y más eficiente.

Hace tan solo diez años los volúmenes de datos eran infinitamente menores a como son ahora. No digamos ya más años atrás. Pero la era IoT y móvil lo ha cambiado todo. Y ha tenido que transformar radicalmente el escenario, partiendo de una perspectiva fundamentada en las bases de datos pero con nuevo aprendizaje desde el mundo del Big Data. Ahí es donde mejor encaja Snowflake, dando a las compañías la posibilidad de hacer algo con esos datos para ganar valor en su negocio. Tenemos la capacidad de integrar los datos desde el backoffice al front end, desde los ERPs y CRMs, teniendo también en cuenta la información no estructurada para ofrecer una mejor experiencia a los clientes, mejores ofertas, etc.

¿Puede concretar cómo impactan esas nuevas tendencias en la gestión del dato?

Los algoritmos obtenidos de modelos de machine learning y de experiencias en Inteligencia Artificial necesitan más datos, pero es un hecho que a más datos, más importantes serán los resultados. Con esto quiero decir que si cuentas con pocos datos, tendrás un análisis poco detallado, pero a más volumen de información los análisis serán



mucho más complejos, difíciles de hacer, y al mismo tiempo más valiosos para hacer predicciones y tomar decisiones basadas en datos. Por eso, será siempre mejor tener más datos, y por tanto, contar con la tecnología capaz de poder llevarlo a cabo.

Snowflake facilita la posibilidad de cargar todos los datos en un único repositorio, para poder tener diferentes aproximaciones al dato. De esta manera, todas las personas trabajan con los mismos datos: se evitan duplicidades y errores. Esa es una de nuestras claras diferenciaciones frente a otras ofertas del mercado. Tampoco hay un lock-in respecto a las diferentes posibilidades de acercarse al dato, o unas herramientas determinadas: cualquier aplicación puede ser conectada a Snowflake: ofrecemos todas las posibilidades y todos los estándares del mercado para incorporar y analizar información. Pero lo verdaderamente importante es que siempre sean los mismos datos, para evitar así los temidos silos, en los que un mismo dato está presente en múltiples plataformas, y múltiples empleados hacen sus propios análisis, obteniendo resultados dispares. Nosotros, además, separamos los recursos de computación del almacenamiento, para eliminar así inconsistencias y al mismo tiempo reducir latencias.

¿En qué compañías encaja Snowflake?

Básicamente en todas. Como tenemos el data warehouse como servicio y apenas necesita configuración, pueden empezar pymes con poco volumen de datos, sin saber todavía hasta dónde llegarán a crecer, para ir demandando más capacidades conforme vayan necesitando. Pero también tenemos grandes empresas que se apoyan en Snowflake para mover sus grandes cargas de datos, incluso las que tienen mucho 'legacy' y desean migrar a la nube. Nuestra plataforma es muy flexible.

¿Cómo funciona Snowflake en la nube y cómo están respondiendo las compañías ante la idea de llevar al cloud su data warehouse?

Snowflake no existiría sin la nube. Es una solución única y exclusivamente cloud, y su arquitectura fue creada específicamente para este entorno. Desde un primer momento, queríamos tener con una solución completamente agnóstica, por lo que la plataforma corre tanto en AWS como en Azure, y próximamente también en Google Cloud. Su arquitectura es completamente independiente del proveedor cloud, pero por supuesto depende de las funcionalidades del proveedor cloud, de su capacidad de almacenamiento y computación. La nube es el futuro de la gestión de los datos.

En la mayor parte de las organizaciones, su 'core' es su negocio, y no las TI. Un data warehouse como servicio elimina la necesidad de tener que contar en plantilla con expertos en esta materia, porque de eso ya nos ocupamos nosotros, junto con las tareas de parcheo, actualizaciones, etc. No hay que preocuparse por cómo se suben los datos o se parametrizan. La principal preocupación que suele haber en torno a la subida de datos a la nube suele ser la seguridad. Pero nosotros lo sabemos, y por eso encriptamos todos los datos antes de ser cargados a la base de datos, de tal manera que los datos ya están encriptados antes de entrar en la nube. Nos aseguramos de cum-

A más datos, más valiosos serán los resultados. Por eso, siempre será mejor tener más datos y contar con la tecnología que permita explotarlos





plir con los protocolos de seguridad estipulados según las normativas locales, y también contamos con numerosas certificaciones que garantizan que cumplimos con la legislación.

¿Cuál es el posicionamiento de Snowflake en Europa?

Llegamos al mercado europeo hace poco más de dos años, con tan solo tres personas en el equipo en Reino Unido. En España abrimos oficinas hace dos meses, y ya tenemos cinco empleados. Nuestro crecimiento ha sido espectacular. Ahora tenemos al menos 60 empleados en Londres, y 150 personas en Amsterdam, y hemos abierto un centro de I+D en Berlín. Empezamos en California y ya hemos dado el salto a muchos otros países, lo que demuestra nuestro gran crecimiento. Seguimos ampliando plantilla en Europa, seguimos extendiéndonos a nuevos países y territorios. Tenemos dos centros de datos, uno en Frankfurt y otro en Dublín, desde los que damos servicio a nuestros clientes europeos, y pronto abriremos otro en los Países Bajos. En Madrid detectamos un gran interés en nuestro data warehouse como servicio, y estamos empezando, pero por supuesto necesitamos seguir dándonos a conocer. Pienso que Snowflake es una compañía muy innovadora y va a cambiar la forma en la que los datos y la analítica deben ser utilizados, y brinda oportunidades a las compañías para que puedan hacer cosas que antes simplemente no podían hacer.

¿Cuánto tiempo se tarda en tener Snowflake en producción?

Unas 24 horas desde que se da de alta en el servicio hasta que se pueden empezar a cargar datos. La carga de datos se hace en función de los datos almacenados, y es una tasa mensual la que se cobra en función del promedio de almacenamiento a lo largo del mes, lo que quiere decir que no se paga por el máximo de datos cargados, sino por la media. Esto es bastante nuevo frente a otras aproximaciones tradicionales en las que aunque hayas ingresado 10 Terabytes de datos un solo día del mes, ya te cobran por esos 10 Teras, a pesar de que el resto del mes apenas hayas introducido datos en el sistema. Aquí se cobra por la media de lo que hayas subido a lo largo del mes, una aproximación mucho más honesta y justa.

¿Qué podemos esperar en un futuro inmediato de la gestión del dato?

Mientras más datos haya disponibles, más útiles serán los resultados obtenidos de su análisis. Llegará el día en el que, igual que cuando haces una búsqueda en Google y te anticipa el resultado, ocurrirá lo mismo con la analítica de datos. Basándose en tus búsquedas y tu historial, el sistema te podría sugerir preguntas sobre qué tipo de análisis quieres hacer. No sé cuando ocurrirá, pero de aquí a un par de años, será perfectamente viable valerse del machine learning para anticiparse a lo que quieres hacer con tus datos. ●