

MUNDO DIGITAL

# 'Big Data' R-evolucionana el manejo de datos

Tal parece que la información siempre tendrá un lugar fundamental en el ámbito de las tecnologías

Foto: Archivo / El Vigía



DARDANE RODRÍGUEZ VERDUGO  
Ensenada, B. C.

A lo largo de la historia la tecnología se ha topado con la necesidad de almacenar datos para su posterior procesamiento. Desde la arquitectura Von Neumann hasta el actual paradigma de la computación en la nube, la información siempre tendrá un lugar fundamental en el ámbito de las tecnologías.

La revolución de los datos se ha dado gracias a la evolución dentro de las telecomunicaciones; la forma en la cual las personas se comunican ha cambiado siendo cada día más fácil compartir cualquier aspecto de la vida y con ello incrementar de manera exponencial la carga de datos en los diferentes servicios existentes en la web.

Con el crecimiento en el manejo de los datos dentro de la red de redes, en la actualidad se ha acuñado el término "Big Data", dicho término se utiliza como sinónimo de este fenómeno.

## MÁS IMPORTANCIA AL ALMACENAMIENTO

Con el auge del "Big Data" se ha redefinido el concepto de dato como tal, dándole más importancia a la parte del almacenamiento, y de cierta manera se ha llevado al olvido por el usuario promedio el tema del manejo de la información, ya que por lo general, cuando se habla de datos, se hace alusión –aunque sea de manera discreta– al manejo de base de datos siendo ésta la forma más conocida y a su vez desconocida dentro de la manipulación de la información.

Para las personas entendidas del tema, escuchar la frase "base de datos" les trae a la mente diagramas, relaciones, entidades y un sinfín de lineamientos para la creación –y acceso– a las mismas. Con esto se da por hecho que todas las bases de datos guardan cierta similitud entre ellas y en parte así es, pues desde SQLserver de Microsoft pasando por PostgreSQL hasta el ya bien conocido MySQL utilizan el mismo modelado y el mismo lenguaje estructurado de consultas (SQL); claro,

cada uno con sus variantes pero la base es la misma.

Existen diversos tipos de bases de datos en el mercado más allá del aceptado modelo relacional, existen bases de datos orientadas a objetos, bases de datos multidimensionales, orientada a grafos, entre otras. No todo es SQL, pero sí una gran parte, es por ello que dicho acrónimo es casi sinónimo de base de datos.

## TECNOLOGÍA DETRÁS DEL NOSQL

Con el crecimiento exponencial de los datos dentro de internet empezaron a surgir nuevas necesidades, las cuales eran solventadas por tecnologías que no eran del tipo relacional y que no utilizaba una sintaxis parecida a SQL, es por ello que desde hace algunos años se ha visto un incremento en el uso de una nueva tecnología denominada NoSQL (Not Only SQL) la cual consiste básicamente en olvidar todo lo que implica modelo clásico de entidad-relación, esto significa que el desarrollador contará con menos trabas en el manejo

de datos, mayor eficiencia en el tratamiento de la información con consultas mucho más sencillas, pues al no utilizar el modelo entidad-relación los "joins", "equi-joins" y similares quedaron totalmente en desuso.

A pesar de no contar con relaciones, estas bases de datos funcionan bastante bien y su uso es bastante sencillo ya que las NoSQL son utilizadas esencialmente para la manipulación de grandes cantidades de datos así como para facilitar la escalabilidad y replicación de los mismos. Esto significa que estas bases de datos pueden llegar a manejar, por ejemplo, los mensajes de una red social, los cuales pueden llegar a los millones de una manera bastante rápida. En caso que el proyecto en desarrollo prospere se podrán "particionar" los datos de tal manera que estos se encontrarán repartidos en varios servidores teniendo un fácil acceso a ellos y no solamente desde un servidor dedicado.

## EL MERCADO DE LOS DATOS

Las bases de datos NoSQL cuentan con una buena base de entusiastas y, sobre todo, proyectos que dan soporte a dicha

tecnología bajo diferentes motores como lo son MongoDB, CouchDB, Apache Cassandra, Dynamo, BigTable entre otros. Cada una de las bases de datos mencionadas cuenta con sus estilos para el manejo de datos, algunos manejan duplas clave/valor, otras utilizan datos tabulares, mientras que hay otras que almacenan la información en documentos accesibles de manera rápida y sencilla.

Las NoSQL no solamente cuentan con una gran base de entusiastas, sino que también cuenta con el respaldo de proyectos de gran envergadura en el mercado, los cuales han cambiado bases de datos relacionales por NoSQL, por ejemplo Facebook, Twitter, Google, Amazon, LinkedIn, Netflix, CERN LHC, Diaspora, Foursquare, Mtv, SourceForge, entre otros.

En cuanto a velocidad, según dictaminan los ingenieros de Facebook, realizar una escritura de 50 Gigabytes toma aproximadamente 0.12 milisegundos (ms), mientras que la lectura de la misma cantidad de información toma 15 ms. Esto en Mysql toma alrededor de 300 ms para realizar cada una de las operaciones.

## SQL O NOSQL, ESE ES EL DILEMA

Aún hoy día existe una gran polémica en cuanto al uso de NoSQL ya que algunas empresas se encuentran en el dilema de utilizar o no utilizar dicha tecnología. El hecho de migrar depende de las necesidades del proyecto, de la misma manera es posible que ambas tecnologías logren apoyarse entre sí, ya que es totalmente válido y factible utilizar SQL y NoSQL en el mismo proyecto.

Es así como el futuro de NoSQL figura prometedor, pero no por ello se declarará la muerte de todas las bases de datos relacionales ni mucho menos. Cada tecnología tiene sus beneficios y sus puntos débiles. No está por demás comentar que la elección de la tecnología se define de acuerdo a las características del proyecto que se desee desarrollar.

\*El autor es profesor del área de Computación en la Universidad Tecnológica de Tijuana, campus Ensenada. dardane.rodriguez@gmail.com