

# Relación entre las bases de datos NoSQL y Big Data

Luis Rubén Artavia and Mario Villalobos

Escuela de Ingeniería,  
Universidad Latinoamericana de Ciencia y Tecnología,  
ULACIT, Urbanización Tournón, 10235-1000  
San José, Costa Rica  
alumno1,alumno2@ulacit.ac.cr  
<http://www.ulacit.ac.cr>

**Abstract.** Los avances tecnológicos crean una propuesta de nuevas alternativas para facilitar y mejorar la vida del ser humano. Tal tipo de adelantos deben encontrar un equilibrio con el medio en donde serán implementados, y estar acorde con los requerimientos y necesidades de quienes los utilizarán. Desde esa perspectiva, este artículo describe los avances en materia de almacenamiento, procesamiento y gestión de grandes cantidades de información mediante el uso de Sistemas de Gestión de Bases de Datos (SGBD) de alto desempeño, como es el caso de las tecnologías NoSQL. De acuerdo con lo anterior, en este trabajo se lleva a cabo un análisis de la situación actual de los SGBD, y su evolución desde el surgimiento de los primeros sistemas de almacenamiento, hasta las nuevas tecnologías de bases de datos disponibles en el mercado, así como su implementación.

**Keywords:** NoSQL, Big Data, Big Data Analytics, Bases de Datos

## 1 Introducción

En la actualidad uno de los principales activos de las empresas es la información, la cual es almacenada por medio de bases de datos relacionales que son administradas por Sistemas de Gestión de Bases de Datos (SGBD). Estos sistemas permiten recuperar la información y optimizan su desempeño, pues reducen los tiempos de respuesta de las consultas. Sin embargo, los requerimientos para gestionar el acceso y análisis de los grandes volúmenes de información (Big Data), que se producen y se encuentran disponibles hoy en día, han vuelto insuficientes las capacidades de gestión de este tipo de bases de datos y sistemas.

El surgimiento de nuevas tecnologías ha hecho posible que las empresas produzcan y compartan información entre fábricas, oficinas y socios de actividades comerciales, distribuidos en diferentes zonas geográficas. De tal manera, la información que es producida por los equipos en las fábricas, los automóviles (desplazamiento y funcionamiento), los sensores (detección de movimiento), cámaras y aplicaciones como Facebook, Twitter o Youtube puede ser accedida en tiempo

real de forma sencilla. Esto ha provocado el crecimiento exponencial de los datos, y por ende, ha originado problemas al gestionar la información almacenada en las bases de datos relacionales.

Esbozado lo anterior, no cabe duda que la aparición de nuevas tecnologías ha mejorado la capacidad de procesamiento y almacenamiento de los equipos de cómputo, por tal razón, se ha optado por mejorar el rendimiento de sus sistemas de hardware y bases de datos. No obstante, el aumento acelerado del volumen de información y los requerimientos de mejores servicios para su procesamiento y análisis, aumentó la presión sobre la comunidad de desarrollo de sistemas de software. Tal situación produjo la aparición de soluciones más robustas, como es el caso de las bases de datos NoSQL (producto de movimientos de grupos de la comunidad de desarrollo).

Las bases de datos NoSQL son estructuras que permiten almacenar y gestionar información de forma escalable, estas proporcionan alto rendimiento y brindan soporte a miles de usuarios concurrentes y a millones de consultas diarias.

Este artículo tiene como objetivo relacionar las bases de datos NoSQL con las Big Data, tomando en cuenta sus principales características y tipos de estructuras. Para tal efecto, en primer lugar se definirá de forma precisa qué es una base de datos NoSQL, posteriormente, se identificarán las principales bases de datos disponibles en el mercado y las que han sido propuestas por la comunidad académica, además se determinarán las características de las diferentes bases de datos, y se realizará una comparación entre éstas considerando su finalidad (según su orientación).

## 2 Antecedentes

Las bases de datos relacionales representan uno de los principales elementos de apoyo en el procesamiento y gestión de la información. El uso de estas permite realizar las tareas de forma rápida y precisa, para brindar respuestas a los eventos que suceden a corto o mediano plazo.

Con el apoyo de este tipo de herramienta se desarrollaron grandes sistemas transaccionales, usando tecnologías como el procesamiento de transacciones en línea (OLTP) y el procesamiento analítico en línea (OLAP), así como grandes almacenes de datos (data warehousing). Estas bases de datos apoyaron el surgimiento de tecnologías avanzadas para la gestión y análisis de la información, tales como la minería de datos (data mining) y sistemas de inteligencia de negocios (BI).

Sin embargo, cada día el análisis y gestión de datos se ha vuelto más crítico, tomando en consideración el crecimiento del volumen y los nuevos tipos de información, circunstancia que ha generado la aparición de nuevas tecnologías. Es decir, los métodos y técnicas que se habían venido utilizando fueron insuficientes. Como consecuencia de lo anterior, el científico y filósofo Thomas Kuhn acuñó un término que mencionó en sus trabajos científicos y que denominó “el paradigma del cambio (paradigm shift)”. Kuhn indica que las ideas innovadoras vienen en ráfagas y que afectan al mundo en formas no lineales (McCreary & Kelly, 2013).

El cumplimiento del paradigma expuesto por Kuhn se manifiesta en el surgimiento de las nuevas necesidades de almacenamiento, gestión y análisis de la información, así como en la aparición de tecnologías y métodos revolucionarios entre los que se cuenta el movimiento NoSQL. Este último influye en la forma como se lleva a cabo el procesamiento de la información y uso de las comunicaciones.

El surgimiento de estas nuevas tecnologías permite el crecimiento horizontal de la infraestructura<sup>1</sup> (en lugar de un crecimiento vertical<sup>2</sup>). Esto brinda la posibilidad de agregar equipos y componentes con gran capacidad de procesamiento y almacenamiento, a bajo costo. El crecimiento horizontal permite aumentar la capacidad de forma proporcional al incremento que se realiza, en contraposición a lo que sucede con el crecimiento vertical.

Una propiedad importante de las herramientas NoSQL es la eliminación de la barrera del almacenamiento tabular de la información, esto hace posible utilizar diferentes tipos de datos de forma flexible, y evita los procesos de transformación o traducción que se requieren para almacenar y recuperar la información de manera relacional.

La implementación de estas capacidades proporciona un alto desempeño y disponibilidad en la gestión de la información, lo que genera como consecuencia que se sacrifiquen las propiedades ACID (Atomic, Consistent, Isolated, Durable) que poseen las bases de datos relacionales. Las características de NoSQL están basadas en los métodos BASE (Basic Availability, Soft state, Eventual consistency).

### 3 ¿Qué es una Base de Datos NoSQL?

El término NoSQL fue acuñado en 1998 por Carlo Strozzi (Strozzi, 2014) para referirse a una base de datos de código abierto que no tenía una interfaz SQL, aunque si seguía el modelo relacional NoSQL que significa Not Only SQL (no solo SQL), y se usa para agrupar sistemas de administración de bases de datos diferentes a los tradicionales (relacionales). Posteriormente, en el año 2009, Eric Evans reutilizó el término para las bases de datos no relacionales, distribuidas y no conformes con las propiedades ACID propias de los sistemas relacionales.

Los sistemas de Bases de Datos NoSQL se consideran de la edad de Internet, y nacen a partir de la necesidad de gestionar grandes cantidades de información en aplicaciones que soportan millones de usuarios diariamente. Los nuevos requerimientos incluyen:

1. Disponibilidad y control total de los datos
2. Tolerancia a las fallas

---

<sup>1</sup> El crecimiento horizontal conlleva el incremento en el número de equipos con el fin de aumentar la capacidad de procesamiento y almacenamiento de la infraestructura que soporta la base de datos.

<sup>2</sup> El crecimiento vertical de la infraestructura implica el aumento de la capacidad de procesamiento y almacenamiento del equipo que se encuentra en funcionamiento.

3. Almacenamiento de petabytes de información distribuida en miles de servidores
4. Alto desempeño sin límite de nodos con escalabilidad horizontal
5. Administración de grandes cantidades de datos estructurados y no estructurados
6. Mecanismos de consulta
7. Gestionar nuevos objetos mediante operaciones efectuadas en tiempo real.

Asimismo, dichos sistemas no requieren esquemas fijos, son fáciles y rápidos de instalar, usan lenguajes no declarativos, ofrecen un alto rendimiento y disponibilidad, y evitan el uso de operaciones join (lo cual mejora el impacto en el rendimiento de la base de datos).

Las principales características de las bases de datos distribuidas no relacionales, propietarias o no, son las siguientes:

- Consistencia eventual: No se implementan mecanismos rígidos de consistencia como los efectuados en las bases de datos relacionales, en los que la confirmación de un cambio implica comunicarlo a todos los nodos que lo repliquen. Esta flexibilidad hace que la consistencia se produzca cuando no se han modificado los datos durante un periodo. Esto se conoce también como BASE (Basically Available Soft-state Eventual Consistency), en contraposición a ACID, su analogía en las bases de datos relacionales.
- Estructura distribuida: Localizan la ubicación de los datos en los diferentes nodos de la arquitectura, por lo general, se utilizan mecanismos de tablas de hash distribuidas <sup>3</sup>, las cuales se implementan de forma diferente según las distintas distribuciones.
- Escalabilidad horizontal: La implementación típica se realiza en muchos nodos con poca capacidad, para no utilizar grandes servidores o centros de procesamiento.
- Tolerancia a fallos y redundancia.

En algunas implementaciones de bases de datos NoSQL, no se utiliza el lenguaje de consultas SQL (por ejemplo, MongoDB usa JSON), pero en otras se implementan y permiten su uso (e.g. BigTable (GQL)) además, se efectúan cambios para mantener su estructura básica.

## 4 ¿NoSQL es Big Data?

El concepto de Big Data fue formulado en un artículo de Doug Laney, cuya definición hacía referencia a sus tres principales características (ver figura 1): volumen, velocidad y variedad (MongoDB, 2014a). De forma adicional, Phillips hace referencia a tres tendencias (ver Figura 2) que son responsables de innovar a nivel de la capa de generación de la información (datos): crecimiento en la generación, crecimientos de usuarios que usan aplicaciones web y la gran demanda en el uso de tecnologías móviles (Phillips, 2012).

<sup>3</sup> Conocidas en inglés como Distributed Hash Table (DHT).

En general, Big Data es utilizado para referirse a grandes y complejas colecciones de datos (estructurados, no estructurados o semi estructurados) que son difíciles de procesar con las herramientas tradicionales de bases de datos y de procesamiento de información (Sherman, 2014)). Como consecuencia su procesamiento requiere de grandes capacidades, lo cual eleva los costos financieros.

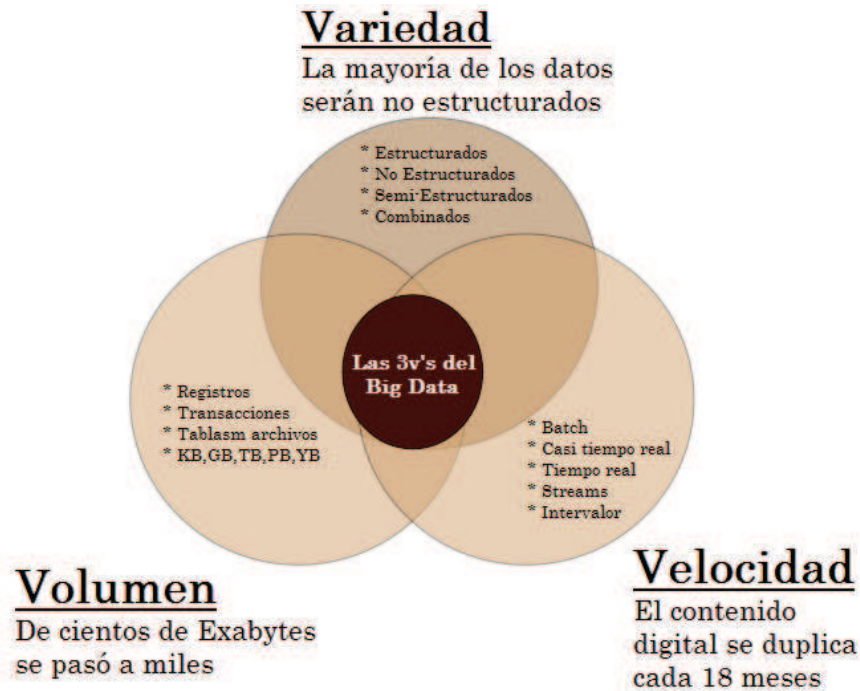
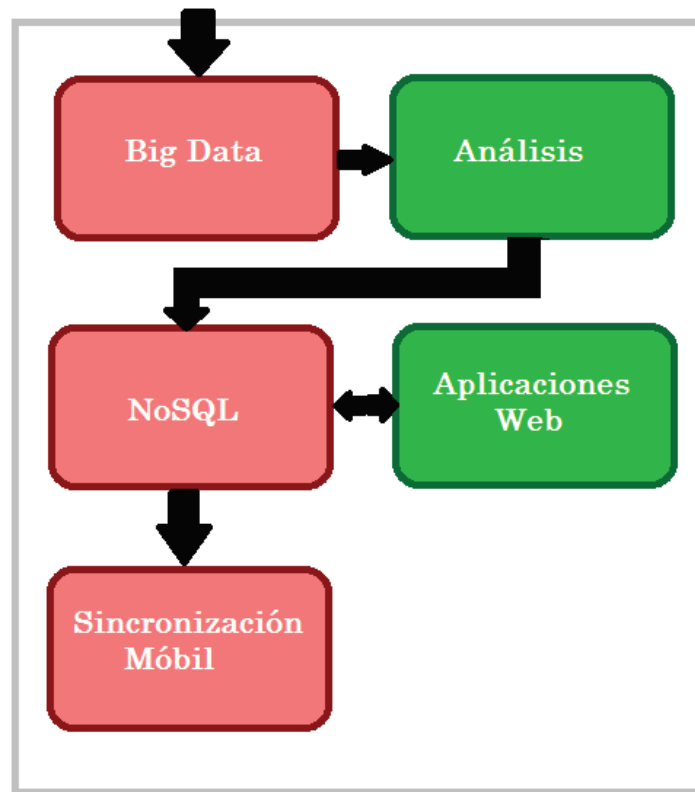


Fig. 1. Características de Big Data

El procesamiento de Big Data tiene como finalidad permitir análisis más sofisticados que produzcan resultados para crear correlaciones entre elementos y ayuden a demostrar la causalidad de las cosas (Sherman, 2014) o sugieran patrones que conduzcan a la producción de conocimiento para la toma de decisiones.

Big Data está dominado por dos clases de tecnologías: sistemas que proporcionan capacidades operacionales en tiempo real, cargas de trabajo interactivas en las que principalmente, los datos se capturan y almacenan; y sistemas que proporcionan capacidades analíticas para retrospectiva, es decir, análisis complejo que puede abarcar la mayoría o totalidad de los datos. Estas clases de tecnología son complementarias y con frecuencia trabajan juntas.



**Fig. 2.** Tendencias de la capa de datos.

Asimismo, cada una de ellas ha impulsado la creación de nuevas tecnologías, tales como las bases de datos NoSQL que están enfocadas en proporcionar servicios para atender solicitudes de información, de forma simultánea con criterios altamente selectivos y respuestas de muy baja latencia.

En este contexto, las características del movimiento NoSQL están basadas en las necesidades de Big Data, es decir, tener la capacidad de almacenar y procesar grandes cantidades de información, así como contar con una infraestructura escalable que permita incrementar la capacidad de almacenamiento y procesamiento, de manera rápida y a bajo costo.

Las tendencias de los últimos años han hecho que los conceptos de NoSQL y Big Data se utilicen como sinónimos, puesto que se consideran complementarios y están estrechamente relacionados. Esto ha provocado debates entre diferentes autores acerca del significado de los términos NoSQL y Big Data (López-de-Ipiña González-de Artaza, 2013) aunque no se refieren al mismo concepto ni grupo de tecnologías (ver Sección 3).

En síntesis, NoSQL hace referencia a la tecnología para almacenar información, y Big Data a la tecnología para analizar dicha información (Gordon, 2013).

## 5 Tipos de Bases de Datos NoSQL

### 5.1 El teorema de Brewer o teorema CAP

Eric Brewer<sup>4</sup> formuló un teorema acerca de las tres dimensiones de los sistemas distribuidos (de almacenamiento de datos en este caso), que son las siguientes (MongoDB, 2014b):

- **Consistencia** (Consistency): Implica que la información permanece coherente y consistente después de cualquier operación sobre los datos, de modo que cualquier usuario que acceda a estos verá la misma información.
- **Disponibilidad** (Availability): Significa que toda la información del sistema de almacenamiento de datos distribuido siempre estará disponible.
- **Tolerancia de las Particiones** (Partition Tolerance): Se refiere a que las diferentes partes del sistema distribuido (los nodos) continuarán funcionando normalmente, aunque la comunicación entre ellos se vea interrumpida o no sea confiable.

Según dicho teorema, no es posible que un sistema de cómputo distribuido cumpla con las tres dimensiones a la vez, por tal motivo hay que elegir dos de ellas; y de los tres grupos resultantes, se generan diferentes tipos de sistemas distribuidos de almacenamiento, a saber:

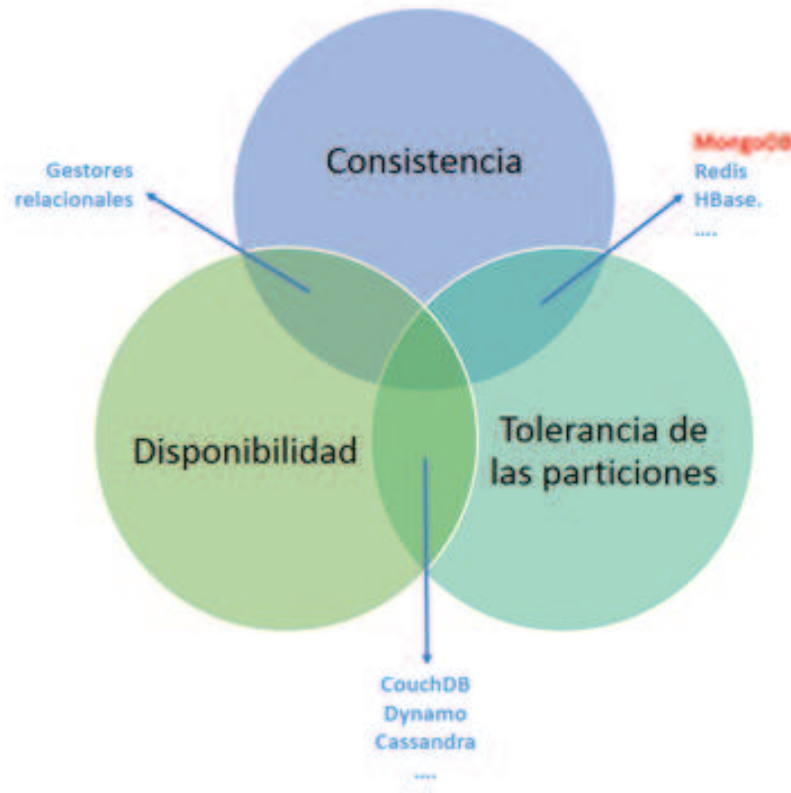
- **C y A**: Estos son sistemas en los que si falla la comunicación entre sus nodos, el conjunto no puede trabajar.
- **C y P**: En estos si algo ocurre, parte de la información no estará disponible, pero seguirán funcionando y la información disponible será consistente.
- **A y P**: Durante un fallo de uno de los nodos (o falta de comunicación) la información estará disponible pero podría ser que no sea consistente.

A partir de esos grupos, podemos clasificar los sistemas de almacenamiento de datos y conocer a qué categoría pertenece cada uno de los sistemas, entre ellos, las Bases de Datos NoSQL, conociendo para cada caso, cuáles características pueden proporcionar, como se muestra en la figura 3.3.

Por ejemplo, MongoDB es un gestor de datos NoSQL distribuido, de tipo documental que por sus características se ubica en la categoría C+P, es decir, ofrece consistencia y tolerancia a las caídas en sus nodos, pero puede ser que alguna información no permanezca disponible.

La situación anterior ha tomado relevancia, debido a que saber la ubicación de un gestor de datos u otro sistema dentro de estos conjuntos, permite tomar decisiones sobre arquitecturas que más tarde pueden ser críticas para nuestra aplicación.

<sup>4</sup> Universidad de California, Berkeley. <http://www.cs.berkeley.edu/~brewer/>, rescatado el 17 de noviembre del 2014.



**Fig. 3.** Teorema CAP, Sistemas de Gestión de Bases de Datos según Dimensión.

## 5.2 Tipos de Bases de Datos NoSQL

Existen diferentes tipos de Sistemas de Administración de Bases de Datos NoSQL, los cuales pueden agruparse en los siguientes (MongoDB, 2014c):

1. Almacenamientos clave-valor, para requerimientos similares a los OLTP (on-line transaction processing) con muchas transacciones sobre pequeñas cantidades de datos, incluyen los sistemas basados en columnas. Son los modelos NoSQL más sencillos de entender, guardan registros (fila o tupla) que contienen una clave y su valor. Cuando se quiere recuperar un dato, se busca por su clave y se recupera el valor. En esta categoría encontramos las siguientes:
  - DynamoDB: Desarrollada por Amazon, es una opción de almacenaje que podemos usar desde los Amazon Web Services. La utilizan el Washington Post y Scopely.
  - Redis: Desarrollada en C y de código abierto, es utilizada por Craigslist y Stack Overflow (a modo de caché).



2. Orientadas a documentos: Son aquellas que gestionan datos semi estructurados, es decir, documentos. Estos datos son almacenados en algún formato estándar como puede ser XML, JSON o BSON.  
Constituyen las bases de datos NoSQL más versátiles. Se pueden utilizar en gran cantidad de proyectos, incluyendo muchos que tradicionalmente funcionarían sobre bases de datos relacionales, a continuación se mencionan algunos de los sistemas más populares:
  - MongoDB: Probablemente la base de datos NoSQL, en la actualidad, es la más famosa.
  - CouchDB: Esta base de datos se orienta a documentos de Apache. Una de sus interesantes características es que los datos son accesibles a través de una API Rest.
3. Orientadas a grafos: Basadas en la teoría de grafos, utilizan nodos y aristas para representar los datos almacenados. Son muy útiles para guardar información en modelos con muchas relaciones, como redes y conexiones sociales; seguidamente entre ellas las más destacadas:
  - Infinite Graph: Escrita en Java y C++ por la compañía Objectivity. Tiene dos modelos de licenciamiento: uno gratuito y otro de pago.
  - Neo4j: Base de datos de código abierto, escrita en Java por la compañía Neo Technology.
4. Orientadas a columnas: Están pensadas para agregar y realizar consultas sobre grandes cantidades de datos. Funcionan de forma parecida a las bases de datos relacionales, pero almacenando columnas de datos en lugar de registros.
  - Cassandra: Utiliza un modelo híbrido, entre orientada a columnas y clave-valor.
  - HBase: Escrita en Java y mantenida por el Proyecto Hadoop de Apache, se utiliza para procesar grandes cantidades de datos.

Al respecto es importante señalar que a medida que se van introduciendo nuevas tecnologías para la gestión de información, surge también la necesidad de innovar en nuevos sistemas de almacenamiento y manejo de esta. Si bien, las bases de datos NoSQL son la respuesta inequívoca, a este creciente mercado de información en el que nos vemos sumergidos, una de sus principales debilidades es su falta de madurez, en contraste, con los sistemas de bases de datos relacionales, los cuales se han mantenido por más de 40 años en el mercado.

### 5.3 Beneficios de los Motores NoSQL

En comparación con las bases de datos relacionales, los sistemas NoSQL son más escalables y ofrecen un rendimiento superior; asimismo, su modelo de datos aborda varios requerimientos para los que el modelo relacional no está diseñado; entre ellos gestión de grandes volúmenes de datos estructurados, semi-estructurados y no estructurados, e implementación de programación orientada a objetos; su uso es fácil y flexible.

En la tabla 1 se expone la comparación entre las bases de datos relaciones o SQL y las bases de datos NoSQL(Prasad & Gohil, 2014).

Característica	Bases de Datos SQL	Bases de Datos NoSQL
Tipo	Solo modelo relacional.	Varios tipos: almacenamiento llave o clave-valor, Orientadas a documentos, orientadas a grafos, orientadas a columnas.
Desarrollo	Desarrolladas en los años 70, como herramientas para almacenamiento de datos.	Surgen en los años 2000's para hacer frente a las limitaciones de bases de datos SQL, en especial sobre la capacidad de escalabilidad, la replicación y almacenamiento de datos no estructurados.
Modelo de Almacenamiento	Registros individuales almacenados como filas en tablas, con columnas que guardan una parte específica sobre ese registro (por ejemplo, "gerente", "fecha contratado", etc.), muy similar a una hoja de cálculo. Datos almacenados en tablas separadas, unidos a través de relaciones; por ejemplo, "oficinas" pueden ser almacenados en una tabla, y "empleados" en otra.	Varía según el tipo de base de datos. Por ejemplo, las bases de datos de almacenamiento key-value (llave-valor) funcionan de manera similar a las bases de datos SQL, pero tienen sólo dos columnas: "llave" y "valor", con información más compleja a veces almacenada dentro de las columnas de "valor". Por su parte, las bases de datos documentales no utilizan el modelo de tabla-y-fila por completo, sino que guardan los datos relevantes juntos en un solo "documento" en formato JSON, XML, u otro tipo, que puede anidar valores jerárquicamente.
Esquema de Almacenamiento	Las estructuras y tipos de datos son creados por adelantado. Para almacenar información de un nuevo elemento, se debe modificar toda la base de datos, por ende, deben deshabilitarse los servicios (bajar la base de datos).	Esquemas de almacenamiento dinámicos. A diferencia de los modelos SQL, los esquemas NoSQL permiten añadir nueva información modificando las estructuras necesarias sin necesidad de deshabilitar la base de datos, asimismo, los datos de tipos diferentes pueden almacenarse juntos según sea necesario.
Escalabilidad	Vertical.	Horizontal.
Modelo de Desarrollo	Mixto: código abierto y cerrado. y cerrado.	Código Abierto (open source).
Manipulación de los Datos	Lenguaje Estructurado de consulta (SQL por sus siglas en inglés), usando comandos como Select, Insert y Update.	A través de API's orientadas a objetos.
Consistencia	Pueden configurarse para implementar altos niveles de consistencia.	Depende del producto, algunos ofrecen niveles fuertes de consistencia, otros no.

**Table 1.** Comparación entre bases de datos SQL y NoSQL

## 6 Casos de uso de las bases de datos NoSQL

En los últimos años, a pesar de lo reciente de la tecnología NoSQL frente a las bases de datos relacionales, el uso de las primeras ha ido en aumento, por parte de empresas prestigiosas. A continuación se incluye una muestra de esos casos, divididos por el tipo de base de datos que utilizan.

### 6.1 MongoDB

MTV Networks: esta compañía es dueña de una gran cantidad de sitios web, con un alto tráfico de información, para la cual era necesario hacer un análisis de los datos que se recolectan. Debido a la diversidad de propiedades, lenguajes, marcos de trabajo y modelos de datos, el uso de un sistema de gestión de bases de datos relacional, vendría a crear complicaciones y afectar el rendimiento, principalmente en las operaciones de lectura.

Luego de determinar las limitaciones de las bases de datos relacionales, además de buscar una solución que permitiera crecer fácilmente y flexibilizar el almacenamiento de datos, consideraron el uso de las bases de datos NoSQL, seleccionando MongoDB como el motor que se ajustaba a sus requerimientos.

Las palabras de Jeff Yemin de MTV fueron las siguientes: *“Desde la perspectiva operacional, [mongoDB] se siente muy parecida a MySQL, por lo que nuestros sistemas y grupos de DBAs están bastante acoplados con la administración”*(MongoDB, 2014d).

Cisco: Empresa dedicada a cambiar la forma de conectarse y comunicarse, en noviembre del 2011 lanzó su red social para empresas, conocida como WebEx Social. Esta plataforma combina cuatro bases que son el núcleo de la Enterprise 2.0: procesos del negocio, contenido, comunicación y la parte social.

Este nuevo producto evolucionó y aumento su crecimiento, e hizo que las complejas consultas SQL utilizadas sobre los esquemas altamente normalizados, empezaran a degradar los tiempos de respuesta; además, existían limitaciones de crecimiento horizontal para soportar las nuevas necesidades, esto sin mencionar las dificultades en la actualización o migración de los esquemas de datos con cada liberación de una nueva versión.

MongoDB ha sido visto como el producto que se ajusta a las características de su herramienta, convirtiéndose en el principal medio de almacenamiento de datos en tiempo real. Biren Gandhi, principal arquitecto para Cisco’s Cloud Collaboration Applications Technology Group (CCATG) mencionó que *“mongoDB es un proyecto de código abierto muy activo, que cuenta con una comunidad vibrante. Este coincide con el conjunto de características de WebEx Social y calza exactamente con nuestras necesidades”*(MongoDB, 2014e).

### 6.2 Hadoop

Amazon WS: Con la evolución que tuvo Amazon.com, pasó de ser una empresa de venta de productos a terceros, a otra para ofrecer servicios en la nube, vio en Hadoop el producto que lo podría apoyar en el almacenamiento de grandes

cantidades de información, implementando con ello el Amazon Elastic MapReduce (Amazon EMR). Este servicio en la web facilita y agiliza el procesamiento de altos volúmenes de datos, a bajo costo (AMAZON, 2014).

Como ejemplo del uso de este nuevo ambiente, tenemos a Yelp.com, dedicada a apoyar a las personas en el proceso de compartir sus opiniones, sobre diferentes tipos de negocios a través de 29 países y más de 120 tipos de mercados. Dentro de los beneficios que obtuvo, está el ahorro en hardware, por un monto aproximado a \$55.000,00 (cincuenta y cinco mil dólares).

### 6.3 Riak

Autoridad Médica de Dinamarca: Esta (DHA por su nombre en Inglés Danish Health and Medicines Authority) creó una tarjeta llamada Common Medicine Card, para administrar de forma centralizada la información concerniente a la salud y la prescripción de tratamientos de todos los pacientes.

Para la implementación de este proyecto era necesario que fuera simple y de bajo costo, pero con capacidad para almacenar gran cantidad de datos, y con un 100% de disponibilidad, para asegurarse de que los ciudadanos recibieran el tratamiento médico correcto.

Por tal motivo, el DHA seleccionó a Riak como el motor de base de datos para almacenar la información de manera centralizada, dado que ofrecía una alta disponibilidad y economía de escala. Como resultado, los fondos destinados para esta institución pueden ser utilizados directamente en el cuidado de los pacientes, en lugar de invertir en grandes infraestructuras para dar soporte (BASHO, 2014).

## 7 Conclusiones

Las Tecnologías de Información y Comunicaciones (TIC) son un servicio a disposición de las organizaciones y el público en general, estas facilitan el desarrollo de las actividades a través de herramientas que permiten automatizar y optimizar procesos de manera segura, lo cual sería difícil y complejo sin las TIC.

Las TIC han evolucionado de acuerdo con las necesidades del hombre, estas ofrecen soluciones para la mayoría de las áreas de trabajo y disciplinas, entre las que se encuentra el almacenamiento y procesamiento de información, por medio de los Sistemas de Gestión de Bases de Datos (SGBD); estos desempeñan una gama de funciones para atender dichas tareas, según las necesidades y capacidades de quien vaya a implementarlos.

Como parte de la actualización de las TIC, los SGBD también han sufrido cambios importantes para mantenerse a la vanguardia, por tal motivo han surgido las Bases de Datos NoSQL, estas operan bajo tecnologías diferentes a los sistemas relacionales (modelos tradicionales con más de 30 años en el mercado), ofreciendo soluciones a las exigencias de almacenamiento y procesamiento de grandes volúmenes de información que podría ser accesada por muchos usuarios a la vez. Este tipo de tareas resultan complejas, y en algunos casos imposibles de gestionar por las herramientas relacionales.

En ese mismo orden de ideas, es importante señalar que las bases de datos NoSQL disponen de diferentes mecanismos para administrar la información según las necesidades de los usuarios. A diferencia de los modelos relacionales, ésta puede almacenarse de forma estructurada o no estructurada según se requiera, asimismo, cada motor NoSQL utiliza su propio lenguaje de consulta y manipulación de datos, en contraparte al SQL (estándar a nivel de los SGBD relacionales).

Otro aspecto importante por considerar es que las herramientas NoSQL tienen un poco más de 10 años en el mercado (aparecieron a inicios del año 2000), situación que para algunos usuarios, podría generar riesgos asociados al servicio de soporte técnico por inopia de personal especializado, situación debida a los recientes que son dichas herramientas en comparación con los sistemas relacionales.

En conclusión, es necesario analizar detenidamente el requerimiento que se deberá satisfacer antes de tomar alguna decisión sobre las TIC por implementar, con la finalidad de disponer de los insumos que permitan seleccionar las herramientas necesarias, para solventar las necesidades que originan su adquisición de forma eficaz y eficiente, conforme a la capacidad de quien vaya a implementarlas.

## References

- AMAZON. (2014, nov).  
pages 12
- BASHO. (2014, nov).  
pages 12
- Gordon, K. (2013). What is big data?. *ITNOW*, 55(3), 12 - 13.  
Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=aci&AN=90069862&lang=es&site=ehost-live> pages 7
- López-de-Ipiña González-de Artaza, D. (2013, jul).  
pages 6
- McCreary, D., & Kelly, A. (2013). *Making sense of nosql*. Manning Publications.  
pages 2
- MongoDB. (2014a, nov).  
pages 4
- MongoDB. (2014b, nov).  
pages 7
- MongoDB. (2014c, nov).  
pages 8
- MongoDB. (2014d, nov).  
pages 11
- MongoDB. (2014e, nov).  
pages 11
- Phillips, J. (2012, March). Big data, nosql and mobile sync – three peas, one pod. *Database Trends and Applications*, 26(1), 27.

Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=aci&AN=73310793&lang=es&site=ehost-live> pages 4

Prasad, A., & Gohil, B. N. (2014). A comparative study of nosql databases. *International Journal of Advanced Research in Computer Science*, 5(5), 170 - 176. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=aci&AN=96929960&lang=es&site=ehost-live> pages 9

Sherman, C. (2014). What's the big deal about big data?. *Online Searcher*, 38(2), 10 - 16. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=95273135&lang=es&site=ehost-live> pages 5

Strozzi, C. (2014, dec).  
pages 3