

# Ingeniería de Servidores Big Data y TPCx-HS

## **Resumen**

En este texto puedes incluir un resumen del documento. Este informa al lector sobre el contenido del texto, indicando el objetivo del mismo y qué se puede aprender de él.

# Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Big Data</b>	<b>3</b>
<b>3. Map Reduce. Hadoop</b>	<b>4</b>
<b>4. Spark y Flink</b>	<b>4</b>
<b>5. Benchmarks: TPCx-HS</b>	<b>4</b>
5.1. Carga de trabajo de TPCx-HS . . . . .	4
5.2. Fases de ejecución del benchmark . . . . .	5
<b>6. Conclusión</b>	<b>6</b>

## 1. Introducción

Desde hace miles de siglos el ser humano ha investigado la manera de almacenar y recopilar información. Durante muchos siglos la escritura y la pintura eran los únicos mecanismos existentes. Posteriormente surgió la fotografía, los discos de vinilo... Sin embargo, poca información seguía ocupando mucho volumen físico. Gracias a los avances tecnológicos de las últimas décadas, hoy en día disponemos dispositivos electrónicos para el almacenamiento de datos binarios. Además, la evolución de estos dispositivos ha sido frenética. IBM comercializó el primer disco duro en 1956. Este constaba solamente de 5 Mega Bytes de capacidad [3] mientras que actualmente podemos utilizar discos duros con más de 1 Tera Byte.

La evolución en la capacidad de cómputo y procesamiento de los computadores también ha sido exponencial. El primer ordenador comercial se presentó en 1951 y se conoce como UNIVAC 1 [7]. Este computador realizaba  $n$  cuentas por segundo. Actualmente hablamos de GHz cuando comparamos la velocidad del procesador de un ordenador. Además, es habitual utilizar ordenadores con 8 o más GB de memoria principal. Se han creado benchmarks con el objetivo de aportar elementos de juicio con los que se permite ver cual o cuáles son las mejores enfoques para conducir a la optimización de los productos.

El benchmarking sirve para obtener informmcción útil que ayude a una organización a mejorar sus procesos [1]. Aún así, el benchmarking es un proceso costoso computacionalmente y además, es un proceso que se ha de reservar para cuestiones de importancia que sean importantes en la organización y no se debe usar en tareas rutinarias [8].

Estas nuevas tecnologías han posibilitado que el almacenamiento de datos sea mucho más sencillo. Podemos guardar multitud de archivos multimedia en un dispositivo de unos centímetros y compartirlos con multitud de usuarios.

El mayor flujo de datos se produce gracias a Internet. Aunque es relativamente joven, se hizo público en 1993, actualmente existen más de mil millones de páginas webs [10]. Además, multitud de dispositivos electrónicos se conectan e interaccionan con Internet (lo que se denomina Internet de las cosas [4]). Los usuarios de estos dispositivos utilizan aplicaciones web y redes sociales, publicando textos y archivos multimedia.

Todo este cúmulo de tecnologías y actividades ha dado lugar a que hoy en día haya más de 10 Zeta Bytes de información almacenados ( $1 \text{ ZB} = 10^{12} \text{ GB}$ ) (ver la Figura 1). De hecho, podemos observar que cada año se generan varios Zeta Bytes de información, el crecimiento es exponencial. Hasta 2003 se habían almacenado en total 5 Exa Bytes ( $10^9$  Giga Bytes) de información. Actualmente, generamos esta cantidad de datos en dos días [9].

Por tanto, existen multitud de aplicaciones donde se tienen que procesar conjuntos de datos masivos. Estos conjuntos de datos requieren nuevas tecnologías y algoritmos que permitan tratarlos eficientemente. El desarrollo de estas nuevas tecnologías y algoritmos es lo que se conoce como Big Data.

En este trabajo presentamos una introducción a Big Data y a las diferentes herramientas software existentes, destacando el papel de la ingeniería de servidores en este contexto.

El resto del texto se organiza como sigue. La Sección 2 contiene una mayor descripción del concepto de Big Data. En la Sección 3 introducimos el paradigma de programación map reduce y la tecnología que lo implementa, denominada Hadoop. En la Sección 4 describimos las nuevas tecnologías que han surgido para cohesionar la filosofía map reduce con el procesamiento iterativo. Estas se denominan Spark y Flink. En la Sección 5 explicamos uno de los benchmarks existentes para las tecnologías Big Data, denominado TPCx-HS. Por último, en la Sección 6 presentamos las conclusiones obtenidas.

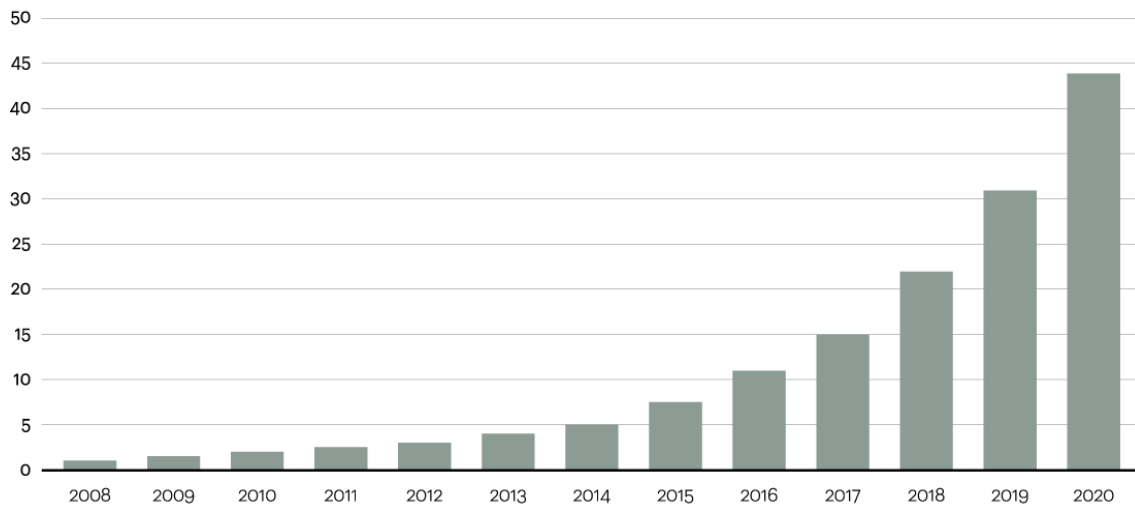


Figura 1: Evolución histórica del número de Zeta Bytes de información almacenados y predicción para los próximos años [2].

## 2. Big Data

A pesar de la evolución de los computadores en todas sus facetas, la cantidad de datos e información a procesar y almacenar crece incluso a mayor velocidad. En muchas ocasiones un ordenador normal no es capaz de tratar tantos Giga Bytes de datos. Estos conjuntos de datos se denominan masivos. Además, en múltiples aplicaciones se necesita aplicar algoritmos sobre los conjuntos de datos recopilados, requiriendo mucho tiempo de cómputo en el caso de que estos sean de gran tamaño.

Big Data engloba el tratamiento de datos masivos desde el punto de vista tecnológico y algorítmico. En palabras de Michael J. Franklin, profesor de informática en la universidad de Berkley [11]:

*“Un problema sobre datos entra en el ámbito de big data cuando la aplicación de las actuales tecnologías no permite al usuario obtener soluciones rápidas, efectivas en costo y de calidad”*

En la literatura especializada se destacan las siguientes características de Big Data, que se denominan las 3 V's de Big Data [9]:

- **Volumen.** El tamaño de los conjuntos de datos a procesar es cada vez mayor, por ejemplo, facebook procesa cada día 500 TB de información. Este volumen de datos requiere tecnologías específicas para que los servidores de altas prestaciones puedan manejar la información con éxito.
- **Velocidad.** Necesitamos herramientas que permitan procesar y analizar conjuntos de datos masivos en poco tiempo. Además, es habitual que el procesamiento de los datos deba ser incluso en tiempo real, esto es, los datos llegan al sistema de forma continua y este debe agregar la información de los mismos.
- **Variedad.** Los datos a tratar provienen de una gran variedad de fuentes. Por tanto, las herramientas Big Data deben permitir procesar a la vez datos de diferentes características y tamaños. Es más, habitualmente encontramos datos de tres tipos: estructurados, semi estructurados y sin estructurar. Los datos estructurados son sencillos de clasificar. Sin embargo, los datos sin estructurar son aleatorios y difíciles de analizar. Por su parte, los datos semi estructurados requieren técnicas avanzadas para poder clasificarlos correctamente.

Algunos autores han extendido la definición hasta utilizar un total de 9 V's: veracidad, valor, viabilidad y visualización entre otras [14].

Los conjuntos de datos contienen conocimiento que necesitamos extraer. Por ejemplo, todas las empresas

almacenan información sobre sus clientes, la actividad y transacciones realizadas. Esta información necesita ser analizada en tiempo real para poder actuar en consecuencia. Aquella empresa que mejor conozca el mercado y actúa rápidamente obtendrá mejores resultados. La ciencia de datos es la rama de la inteligencia artificial que se encarga de tratar y extraer conocimiento de los datos (referencia). Se han desarrollado múltiples algoritmos para ello (referencia machine learning), que permiten resultados tan impactantes como el reconocimiento de voz o los sistemas de recomendación. Big Data Analytics [5].

### 3. Map Reduce. Hadoop

Por tanto, necesitamos recurrir a servidores de altas prestaciones y procesamiento distribuido para poder aplicar estos algoritmos.

Los servidores de altas prestaciones (alguna descripción). Se han desarrollado herramientas de cómputo en paralelo y distribuido, como OPEN MP y MPI (referencias), que permiten implementar algoritmos distribuidos sobre estos. Sin embargo, estas implementaciones dependen del servidor y son a bajo nivel. Una determinada implementación sobre un esquema hardware puede funcionar bien en determinado momento pero al año tendrá que ser capaz de trabajar con el doble de datos. Esto probablemente suponga la necesidad de ampliar el hardware y rehacer la implementación. Necesitamos pues nuevos paradigma de programación que permitan abstraer el desarrollo de software para plataformas distribuidas del hardware y proporcionen capacidad para el tratamiento de datos masivos.

### 4. Spark y Flink

### 5. Benchmarks: TPCx-HS

Una técnica utilizada para ver qué sistema nos da más prestaciones son los benchmarks, uno de ellos es TPC. TPC es una organización sin ánimo de lucro que estudia el proceso de transacción y los benchmarks para las bases de datos. El término transacción se suele atribuir a aspectos bancarios, pero si pensamos en este concepto como una función que tienen los ordenadores, una transacción puede ser un conjunto de operaciones que incluyen lectura y escritura en disco, llamadas a funciones del sistema operativo, o cualquier forma de transferencia de datos de un sistema a otro. Este es el ámbito en el que se mueve TPC, que produce benchmarks que miden el proceso de transacción y el rendimiento de las bases de datos en términos de: dados un sistema y una base de datos, ¿cuántas transacciones pueden hacer por unidad de tiempo?[12]

TPCx-HS se desarrolló para proveer de un rendimiento fiable, rendimiento relación precio, disponibilidad y, opcionalmente, datos de consumo de energía de los sistemas de Big Data. TPCx-HS fue el primer benchmark objetivo que permitía medir tanto hardware como software, como el Hadoop Runtime. [13]

#### 5.1. Carga de trabajo de TPCx-HS

La carga de trabajo de TPCx-HS consiste en los siguientes módulos:

- HSGen: es un programa que genera los datos según factor de escala, que suele estar entre un 1TB y 10000TB(se denota TB como terabytes).
- HSDataCheck: es un programa que comprueba el cumplimiento del conjunto de datos.
- HSSort: es un programa que ordena los datos según un orden total.

- HSValidate: es un programa que valida la salida, es decir, los resultados obtenidos.

## 5.2. Fases de ejecución del benchmark

Una ejecución válida consiste en cinco fases separadas que corren secuencialmente. Estas fases no se solapan en su ejecución, es decir, el comienzo de la fase 2 no puede darse hasta que la fase 1 esté completa. Para que comience cada fase se necesita de un script llamado ¡TPCx-HS-master!, que es el que inicia cada fase y que puede ser ejecutado desde cualquier nodo del sistema que se está bajo test.

Las fases de ejecución son las siguientes:

- Fase 1: se generan los datos de entrada usando HSGen. Se han de copiar en un soporte duradero y hacer la copia replicada en tres discos (llamado "*3-ways replication*"), que suelen ser el principal, el secundario y uno de respaldo[6].
- Fase 2: se verifica la validez del conjunto de datos usando HSDataCheck. El programa sirve para verificar la cardinalidad, tamaño y el factor de réplica de los datos generados. Si el programa reporta un fallo, entonces la ejecución no es válida y se deberá volver a empezar.
- Fase 3: se lanza el programa HSSort con el que se ordena los datos de entrada. Esta fase muestra los datos de entrada y los datos de salida (los datos ya ordenados). Al igual que en la fase 1, se han de copiar los datos en un soporte duradero y hacer el "*3-ways replication*".
- Fase 4: se comprueba la viabilidad del conjunto de datos usando HSDataCheck. El programa comprueba la cardinalidad de los datos, tamaño y el factor de replicación de los datos ordenados. Si el programa reporta un fallo, entonces la ejecución no es válida y se deberá volver a empezar.
- Fase 5: validación de los datos con HSValidate. Como su nombre indica, HSValidate comprueba que los datos de salida sean correctos y si reporta el fallo consisten en que el HSSort no generó el correcto orden de salida, la ejecución se considerará inválida.

El benchmark consiste en dos ejecuciones (ver la Figura 2) y cada vez que se ejecuta una fase se ha de especificar el tiempo que ha llevado la ejecución. Entre la primera ejecución de las cinco fases y la segunda, hay una fase intermedia que sirve para limpiar el sistema, conocida como "*file system cleanup*" y, lógicamente, no se permite ninguna actividad durante dicha fase. Una vez realizadas las dos ejecuciones se obtiene el tiempo total de ejecución que servirá para calcular datos en la fase de medición.

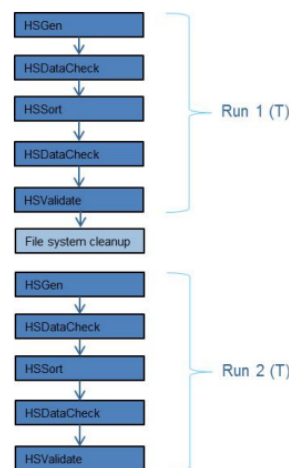


Figura 2: Gráfico de las ejecuciones de TPCx-HS

Como consideraciones finales en la ejecución, el sistema bajo test no puede ser reconfigurado o cambiado durante o entre cualquiera de las fases de la ejecución ni tampoco entre la primera y la segunda ejecución. Cualquier cambio que se haga en el sistema se deberá realizar antes del comienzo de la fase 1 de la primera ejecución.

## 6. Conclusión

## Referencias

- [1] Confederación Granadina de Empresarios. *¿Qué es el Benchmarking*. URL: <http://www.cge.es/portalcge/tecnologia/innovacion/4111benchmarking.aspx>.
- [2] Hugo Evans y et. al. *Big Data and the Creative Destruction of Today's Business Models*. URL: [http://www.atkearney.es/paper/-/asset\\_publisher/dVxv4Hz2h8bS/content/big-data-and-the-creative-destruction-of-today-s-business-models/10192](http://www.atkearney.es/paper/-/asset_publisher/dVxv4Hz2h8bS/content/big-data-and-the-creative-destruction-of-today-s-business-models/10192).
- [3] Rex Farrance. «Timeline: 50 Years of Hard Drives». En: *PCWorld* (2016). URL: <http://www.pcworld.com/article/127105/article.html>.
- [4] F. Herrera. *Inteligencia artificial, inteligencia computacional y Big Data*. Servicio de Publicaciones, Universidad de Jaén, 2014.
- [5] Karthik Kambatla y col. «Trends in big data analytics». En: *Journal of Parallel and Distributed Computing* (2014), págs. 2561-2573.
- [6] Linbit. *Three-way replication*. URL: <https://www.drbd.org/en/doc/users-guide-83/s-three-way-repl>.
- [7] Computer History Museum. *Timeline of Computer History*. URL: <http://www.computerhistory.org/timeline/1951/>.
- [8] Universidad Técnica José Peralta. *Ventajas, desventajas y causas posibles de fracasos del benchmarking*. URL: <http://www.buenastareas.com/ensayos/Ventajas-Desventajas-y-Causas-Posibles-De/3531988.html>.
- [9] Seref Sagiroglu y Duygu Sinanc. «Big data: A review». En: *International Conference on. IEEE* (2013), págs. 42-47.
- [10] Internet live stats. *Total number of Websites*. URL: <http://www.internetlivestats.com/total-number-of-websites/>.
- [11] Steve Todd. *AMPed At UC Berkeley*. URL: [http://stevetodd.typepad.com/my\\_weblog/2011/08/amped-at-uc-berkeley.html](http://stevetodd.typepad.com/my_weblog/2011/08/amped-at-uc-berkeley.html).
- [12] TPC. *About TPC*. URL: <http://www.tpc.org/information/about/abouttpc.asp>.
- [13] TPC. *Transaction Processing Performance Council (TPC) Launches TPCx-HS, the First Vendor-Neutral, Industry Standard Big Data Benchmark*. URL: [http://www.tpc.org/information/other/tpcx-hs\\_press\\_release\\_final.pdf](http://www.tpc.org/information/other/tpcx-hs_press_release_final.pdf).
- [14] P. Zikopoulos y C. Eaton. *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media, 2011.