

Análisis de PostgreSQL, Microsoft Access y VoltDB

Jesús Villacampa

29 de abril de 2020

Índice

1. PostgreSQL	1
2. Microsoft Access	3
3. VoltDB	4
Referencias	5

1. PostgreSQL

PostgreSQL, también llamado *Postgres*, es un sistema de gestión de bases de datos relacional orientado a objetos y de código abierto. Como muchos otros proyectos de código abierto, el desarrollo de *PostgreSQL* no es manejado por una empresa o persona, sino que es dirigido por una comunidad de desarrolladores que trabajan de forma desinteresada, altruista, libre o apoyados por organizaciones comerciales. Dicha comunidad es denominada *PGDG*¹. [1]

PostgreSQL está publicado bajo *PostgreSQL License*, una licencia *Open Source*, similar a *BSD*² o *MIT*³. [2]

Uno de los puntos fuertes de *Postgres* es el cumplimiento de las propiedades ACID⁴ y su gestión de la concurrencia. Aseguran que una lectura nunca bloquea escrituras y viceversa. Esto se consigue a través de un mecanismo llamado *MVCC*⁵. Cada transacción en *Postgres* obtiene una ID de transacción llamada *XID*. Esto incluye transacciones de una sola declaración, como insertar, actualizar o eliminar, así como envolver explícitamente un grupo de declaraciones juntas a través de *BEGIN-COMMIT*. Cuando comienza una transacción, *Postgres* incrementa un *XID* y lo asigna a la transacción actual. [3]

Independientemente de la plataforma y la arquitectura que usemos, *PostgreSQL* está disponible para los diferentes SO, *Unix*, *Linux* y *Windows*, en 32 y 64 bits. Ésto hace de *PostgreSQL* un sistema multiplataforma y también hace que sea más rentable con instalaciones a gran escala. Además tiene más de 20 años de desarrollo activo y en constante mejora. No se han presentado nunca caídas de la base de datos. Ésto es debido a su capacidad de establecer un entorno de alta disponibilidad gracias a *Hot-Standby*⁶. [5]

Postgres está diseñado para ambientes de alto volumen, lo que hace que la velocidad de respuesta en inserciones o actualizaciones pueda parecer lenta en comparación con bases de datos de pequeño tamaño. [6]

Otro aspecto muy importante a tener en cuenta a la hora de elegir *SGBD* es el soporte que proporcionan. En el caso de *PostgreSQL* no cuenta con soporte en línea o telefónico. Sin embargo, cuenta con foros oficiales donde usuarios exponen sus dudas y son los propios usuarios de la comunidad los que responden. Cabe resaltar que esta comunidad de usuarios es una de las más activas del mercado. Además, también ofrece soporte empresarial como *EnterpriseDB*⁷ o *TodoPostgreSQL*⁸. [6]

Postgres es uno de los gestores con más soporte y recorrido en el mercado, y por tanto, cuenta con una herramienta administración llamada *pgAdmin* (desde 2016 en la versión *pgAdmin4*), que facilita la gestión y administración de bases de datos ya sea mediante instrucciones SQL o con ayuda de un entorno gráfico. Permite acceder a todas las funcionalidades de la base de datos; consulta, manipulación y gestión de datos, incluso opciones avanzadas como manipulación del motor de replicación *Slony-I*⁹. A parte de este entorno de escritorio, también existen otras herramientas como *psql* (cliente de consola)

¹PostgreSQL Global Development Group

²https://es.wikipedia.org/wiki/Licencia_BSD

³https://es.wikipedia.org/wiki/Licencia_MIT

⁴Atomicity, Consistency, Isolation y Durability (Atomicidad, Consistencia, Aislamiento y Durabilidad en castellano)

⁵Multi Version Concurrency Control

⁶Hot-Standby nos permite que los clientes puedan realizar consultas de solo lectura mientras que los servidores están en modo de recuperación o espera. Así podemos hacer tareas de mantenimiento o recuperación sin bloquear completamente el sistema [4]

⁷<https://www.enterprisedb.com>

⁸<https://todopostgresql.com/>

⁹Slony-I es un sistema asíncrono de replicación maestro-esclavo para PostgreSQL, que proporciona soporte para conexión en cascada y conmutación por error. Asíncrono significa que cuando una transacción de base de datos se ha comprometido con el servidor maestro, todavía no se garantiza que esté disponible en esclavos. La conexión en cascada significa que las réplicas se pueden crear (y actualizar) a través de otras réplicas, es decir, no necesitan conectarse directamente al maestro. [7]

o *PhpPgAdmin* (entorno web).[8]

2. Microsoft Access

Microsoft Access es un sistema de gestión de bases de datos incluido en el paquete ofimático denominado *Microsoft Office*, sucesor de *Embedded Basic*. *Access* es un gestor de datos que utiliza los conceptos de bases de datos relacionales y pueden manejarse por medio de consultas e informes. Está adaptado para recopilar datos de otras utilidades como *Excel* o *SharePoint*. A partir del 22 de septiembre de 2015 la versión más nueva es *Access 2016*. [9]

Los requisitos para la utilización de este gestor son un sistema operativo *Windows 10*, un procesador de 2 núcleos a 1'6GHz, una memoria RAM de 2GB, un espacio libre en disco duro de 4 GB y una cuenta Microsoft [10]. El precio para adquirir una licencia para 1 PC de *Access* es de 135€. [11]

Las bases de datos *access* son útiles para pequeñas y medianas empresas o departamentos individuales, ya que cuentan con una capacidad limitada. Cualquier sector cuyo uso vaya más allá de las 2 GB descubrirá sus limitaciones [12]. Para evitar esta limitación de tamaño, se puede crear vínculos a tablas de otras bases de datos de *Access*. Puede crear vínculos a tablas de varios archivos de base de datos, cada uno de los cuales puede tener hasta 2 gigabytes de tamaño. [13]

Microsoft Access ofrece consultas parametrizadas ¹⁰. Estas consultas y tablas *Access* se pueden referenciar desde otros programas como *VB6* ¹¹ y *.NET* a través de *DAO* o *ADO*. [15]

También está el formato de base de datos *Jet* ¹², que puede contener la aplicación y los datos en un archivo. Esto hace muy conveniente para distribuir la aplicación completa a otro usuario, que puede ejecutarla en entornos desconectados. [15]

El límite técnico de las bases de datos *access* es de 255 usuarios al mismo tiempo, pero el límite real es entre 10 y 80 usuarios dependiendo del tipo aplicación. [12]

¹⁰Se denominan consultas paramétricas a aquellas consultas, normalmente de selección, que bajo un campo por el que se desea establecer un criterio de filtrado, en vez de implicar a valores constantes que hacen que dicha consulta al ser ejecutada siempre realice "lo mismo", se implique a campos (especificados entre corchetes []) a los cuales se les asignan valores diferentes para cada ejecución de la consulta (normalmente serán solicitados al usuario mediante una ventana en pantalla).[14]

¹¹Visual Basic 6.0

¹²Jet es un motor de bases de datos que permitiría mantener los archivos .mdb a través de ODBC [16]

3. VoltDB

VoltDB es una base de datos en memoria diseñada por Michael Stonebraker, Sam Madden y Daniel Abadi. Es un *RDBMS* ¹³ compatible con ACID que utiliza una arquitectura de nada compartido ¹⁴. Incluye tanto ediciones empresariales como comunitarias. La edición comunitaria está licenciada bajo la *GNU Affero General Public License*. [18]

VoltDB es una base de datos relacional *NewSQL* ¹⁵ que admite acceso *SQL* desde procedimientos almacenados Java precompilados. La unidad de transacción es el procedimiento almacenado, escrito en Java intercalado con *SQL*. *VoltDB* se basa en la partición horizontal hasta el subproceso de hardware individual para escalar, k-safety (replicación síncrona) para proporcionar una alta disponibilidad y una combinación de instantáneas continuas y registro de comandos para una mayor durabilidad (recuperación de fallos). [18]

Una gran desventaja de este *SGBD* es que solo se ejecuta en *Linux* y *Mac*, y por tanto para su utilización en *Windows* haría falta un despliegue en *contenedores docker* o la utilización de máquinas virtuales. [20]

El punto clave de *VoltDB* es su velocidad. Puede procesar hasta 45 veces más transacciones por segundo que otros gestores tradicionales como *MySQL*, *Oracle* y *PostgreSQL*. *VoltDB* procesó 53.000 transacciones por segundo (TPS) frente a las 1.155 de otros *DBMS* sobre hardware idéntico. [21] Debido a esto, se debe ejecutar en servidores ricos en memoria multi-core y aunque esté disponible de manera gratuita bajo la licencia *GPL* ofrecen un precio de suscripción anual es de 15 dólares por año para una configuración de cuatro servidores.

VoltDB permite operacionalizar *modelos ML*. *VoltDB* convierte automáticamente el *modelo ML* en un proceso ejecutable como una función definida por el usuario (UDF), que puede implementarse en un entorno de producción. El modelo integrado ingiere / entrena continuamente datos nuevos / históricos para mantenerse actualizado y relevante. El ML en la base de datos le permite tomar decisiones inteligentes y procesables en tiempo real con una latencia muy baja en la transmisión de datos. [22]

VoltDB se ejecuta en un entorno distribuido en contenedores. Hasta ahora, orquestar bases de datos SQL en *Kubernetes* ¹⁶ había sido muy desafiante, ya que los sistemas de bases de datos operacionales tienen estado y no pueden ser acelerados o reducidos en cualquier momento. Los desarrollos recientes aseguran la orquestación de contenedores con *Kubernetes*, convirtiendo muchas tareas tediosas y complejas en algo tan simple como un archivo de configuración declarativo, y permitiendo una implementación continua y frecuente. [22]

¹³Sistema de gestión de bases de datos relacionales

¹⁴Una arquitectura de nada compartido es una arquitectura de computación distribuida en la que cada solicitud de actualización es satisfecha por un solo nodo (procesador / memoria / unidad de almacenamiento). La intención es eliminar la contención entre nodos. Los nodos no comparten (acceso independiente) memoria o almacenamiento [17]

¹⁵NewSQL es una clase de sistemas de gestión de bases de datos relacionales que buscan proporcionar la escalabilidad de los sistemas NoSQL para cargas de trabajo de procesamiento de transacciones en línea (OLTP) mientras mantienen las garantías ACID de un sistema de base de datos tradicional. [19]

¹⁶Kubernetes es un sistema de código libre para la automatización del despliegue, ajuste de escala y manejo de aplicaciones en contenedores que fue originalmente diseñado por Google y donado a la Cloud Native Computing Foundation. Soporta diferentes entornos para la ejecución de contenedores, incluido Docker. [?]

Referencias

- [1] "Postgresql." <https://es.wikipedia.org/wiki/PostgreSQL>. Consultado el 28/04/2020.
- [2] "Postgresql license." <https://www.postgresql.org/about/licence>. Consultado el 28/04/2020.
- [3] "Mvcc." <https://devcenter.heroku.com/articles/postgresql-concurrency>. Consultado el 28/04/2020.
- [4] "Hot-standby." <https://todopostgresql.com/tecnica-hot-standby>. Consultado el 28/04/2020.
- [5] "Alta disponibilidad." <https://todopostgresql.com/alta-disponibilidad-postgresql/>. Consultado el 28/04/2020.
- [6] "Ventajas y desventajas." <https://todopostgresql.com/ventajas-y-desventajas-de-postgresql/>. Consultado el 28/04/2020.
- [7] "Sloni-i." <https://en.wikipedia.org/wiki/Slony-I>. Consultado el 28/04/2020.
- [8] "Pgadmin." <https://www.pgadmin.org/>. Consultado el 28/04/2020.
- [9] "Microsoft access." https://es.wikipedia.org/wiki/Microsoft_Access. Consultado el 28/04/2020.
- [10] "Microsoft access requisitos." https://www.microsoft.com/es-es/p/access/cfq7ttc0k7q8?=&ef_id=CjwKCAjwqJ_1BRBZEiwAv73uwNScKHTjdgEy6LznNrbv_YoYDtv_cQXDIQNLEgF7gcGVZtUZCPmVRoCj58QAvD_BwE%3aG%3as&OCID=AID2000750_SEM_xUz8JKCM&MarinID=sxUz8JKCM%7c431595526780%7cms+access%7ce%7cc%7c%7c62233861395%7ckwd-26962656&lnkd=Google_0365SMB_App&gclid=CjwKCAjwqJ_1BRBZEiwAv73uwNScKHTjdgEy6LznNrbv_YoYDtv_cQXDIQNLEgF7gcGVZtUZCPmVRoCj58QAvD_BwE&activetab=pivot:regionofsystemrequirementstab. Consultado el 28/04/2020.
- [11] "Microsoft access licencia." https://www.microsoft.com/es-es/p/access/cfq7ttc0k7q8?=&ef_id=CjwKCAjwqJ_1BRBZEiwAv73uwNScKHTjdgEy6LznNrbv_YoYDtv_cQXDIQNLEgF7gcGVZtUZCPmVRoCj58QAvD_BwE%3aG%3as&OCID=AID2000750_SEM_xUz8JKCM&MarinID=sxUz8JKCM%7c431595526780%7cms+access%7ce%7cc%7c%7c62233861395%7ckwd-26962656&lnkd=Google_0365SMB_App&gclid=CjwKCAjwqJ_1BRBZEiwAv73uwNScKHTjdgEy6LznNrbv_YoYDtv_cQXDIQNLEgF7gcGVZtUZCPmVRoCj58QAvD_BwE&activetab=pivot:overviewtab. Consultado el 28/04/2020.
- [12] "Capacidad access." <https://www.informaticaparatunegocio.com/blog/ventajas-desventajas-las-bases-datos-access/>. Consultado el 29/04/2020.
- [13] "Especificaciones access." <https://support.office.com/es-es/article/especificaciones-de-access-0cf3c66f-9cf2-4e32-9568-98c1025bb47c>. Consultado el 29/04/2020.
- [14] "Consultas parametrizadas." https://www.adrformacion.com/knowledge/ofimatica/consultas_parametricas_en_access.html. Consultado el 29/04/2020.
- [15] "Especificaciones access." <https://www.tecnologias-informacion.com/access.html>. Consultado el 29/04/2020.
- [16] "Motor jet." https://en.wikipedia.org/wiki/Microsoft_Jet_Database_Engine. Consultado el 29/04/2020.
- [17] "Arquitectura de nada compartido." https://en.wikipedia.org/wiki/Shared-nothing_architecture. Consultado el 29/04/2020.

- [18] “Voltdb.” <https://en.wikipedia.org/wiki/VoltDB>. Consultado el 29/04/2020.
- [19] “NewsqL.” <https://en.wikipedia.org/wiki/NewSQL>. Consultado el 29/04/2020.
- [20] “Desventajas y ventajas voltdb.” https://nanopdf.com/download/principales-sgbd_pdf. Consultado el 29/04/2020.
- [21] “Velocidad de voltdb.” <https://itblogpress.wordpress.com/2014/02/25/voltdb-bases-de-datos-45-veces-mas-rapidas-que-los-tradicionales-sistemas-mysql-y-oracle/> Consultado el 29/04/2020.
- [22] “Voltdb.” <https://www.voltdb.com/why-voltdb>. Consultado el 29/04/2020.