

ENSAYO POSTGRES-MYSQL

Un mundo cambiante, Postgres se queda atrás.



Por:

Edwin Rodriguez Colorado

Edilber Gomez

Wilfer Berrio

Juan David Zapata

Materia:

Administración de bases de datos

Profesor:

Andres Martínez

18 de noviembre de 2017

itm institución universitaria

Facultad de ingenierías

En un mundo cambiante, donde la tecnología está evolucionando a cada segundo y donde cada día más personas usan las herramientas de software, se debe almacenar la información de manera muy recursiva, segura y ágil, para garantizar que la información siempre se va a encontrar disponible para cualquier fin, pero a su vez también va a estar lo suficientemente segura, como para que no sea vulnerada por los ladrones de información.

Uber es una plataforma que ofrece el software como servicio, para que las personas puedan usar sus vehículos como medio de transporte para otras personas o para conseguir un transporte de manera, cómoda, fácil y ágil, esta compañía de software ha logrado inclinar cada día mas a las personas a usar su plataforma y gracias al logro en la consecución de las metas también ha adquirido más retos en su esquema de desarrollo y almacenamiento de información, por lo que han empezado a descubrir problemas que inicialmente no habían notado o no habían considerado traumáticos debido a que la cantidad de información que manejaban no tenía un volumen considerable como para encontrar las imitaciones del motor de base de datos que manejaban, Uber inicialmente empezó a trabajar con Postgres y gracias a esto logró crear el modelo de datos que necesitaban, modelos que incluso hoy se han sostenido, así como también hay otros modelos que han tenido que ser migrados basados en modelos de otras estructuras para poder seguir sosteniendo el crecimiento continuo de su base de datos.

Postgres es un DBE (DATABASE ENGINE) motor de base de datos que ha ido perdiendo fuerza en el tiempo, ya muchas otras grandes compañías han migrado este motor, pero ese tema no lo hablaremos en este ensayo, lo que sí vamos a hablar es de cuáles fueron las razones que dio Uber para migrar sus bases de datos, a continuación, enumeramos algunas:

* Arquitectura ineficiente para escribir:

La arquitectura de Postgres se basa en Tuplas, las cuales en una base de datos con muchos índices secundarios genera demasiada ineficiencia, ya que para actualizar una fila debe actualizar hasta índices que no hay necesidad de editar, Por ejemplo, si tenemos una tabla con una docena de índices definidos en ella, una actualización de un campo que solo está cubierto por un único índice debe propagarse a los 12 índices para reflejar el ctid de la nueva fila.

* Replicación de información ineficiente

El problema de amplificación de escritura también se traduce naturalmente en la capa de replicación porque la replicación se produce en el nivel de los cambios en el disco, en lugar de replicar un pequeño registro lógico.

* Problemas con corrupción de tablas

Durante una promoción de base de datos maestra de rutina para aumentar la capacidad de la base de datos, nos encontramos con un error de Postgres 9.2. Las réplicas siguieron los cambios de la línea de tiempo de forma incorrecta, haciendo que algunos de ellos apliquen incorrectamente algunos registros WAL. Debido a este error, algunos registros que deberían haber sido marcados como inactivos por el mecanismo de control de versiones no estaban realmente marcados como inactivos. Este problema es bastante irritante ya que no es fácil encontrar a cuantas filas afectó este problema.

* Soporte pobre de réplica MVCC

Postgres no tiene una verdadera réplica de soporte MVCC. El hecho de que las réplicas apliquen las actualizaciones de WAL da como resultado que tengan una copia de los datos en disco idéntica a la maestra en cualquier punto dado en el tiempo. Este diseño plantea un problema para Uber.

* Dificultad para actualizar a lanzamientos más nuevos

Debido a que los registros de replicación funcionan en el nivel físico, no es posible replicar datos entre diferentes versiones de disponibilidad general de Postgres. Una base de datos maestra que ejecuta Postgres 9.3 no puede replicarse en una réplica que ejecuta Postgres 9.2, ni un maestro que ejecuta 9.2 se puede replicar en una réplica que ejecuta Postgres 9.3. Se comienza con Postgres 9.1 y se completa con éxito el proceso de actualización para pasar a Postgres 9.2. Sin embargo, el proceso tomó tantas horas que no se pudo hacer el proceso nuevamente. Para cuando salió Postgres 9.3, el crecimiento de Uber aumentó sustancialmente el conjunto de datos, por lo que la actualización hubiera sido incluso más larga. Por esta razón, las instancias heredadas de Postgres ejecutan Postgres 9.2 hasta el día de hoy, a pesar de que la versión actual de Postgres GA es 9.5.

Todas estas grandes problemáticas llevaron a Uber a pensar en cambiar el motor de base de datos migrando a MySql y a realizar algunas tareas No Sql con software como Casandra, MySql resulta ser una gran herramienta e incluso también es más económica que Postgres.

Luego de una evaluación Uber adquiere el motor de base de datos MySql InnoDB que no solo es usado por Uber, ya que este es lejos el DBE más popular.

La diferencia arquitectónica más importante es que, mientras Postgres correlaciona directamente los registros de índice con las ubicaciones en disco, InnoDB mantiene una estructura secundaria. En lugar de mantener un puntero a la ubicación de la fila en el disco (como hace el ctid en Postgres), los registros del índice secundario de InnoDB tienen un puntero al valor de la clave primaria. Por lo tanto, un índice secundario en MySQL asocia claves de índice con claves principales.

MySql usa un modo de replicación llamado replication Binary Log el cual es considerablemente más compacto que el WAL de Postgres.

MySql fue probado en todos los puntos donde Postgres falló y mostro que puede con grandes volúmenes, mostró que puede manejar información considerable y delicada y ante todo que permite que los upgrade se lleven a cabo, teniendo en cuenta los negocios dinámicos de hoy sin afectar las operaciones diarias.

En conclusión, Postgres es una muestra de la velocidad en que el mundo va cambiando y que, si no se está creciendo y transformando constantemente tu negocio adaptándose a la estructura del día a día, el negocio va a decaer sin importar lo bueno que haya sido y si ha logrado hacer grandes aportes a la humanidad.